THE

INTERNATIONAL SERIES

OF

MONOGRAPHS ON PHYSICS

GENERAL EDITORS

R. H. FOWLER AND P. KAPITZA

THE INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS

GENERAL EDITORS: R. H. Fowler and P. Kapitza

Already Published

WAVE MECHANICS. ELEMENTARY THEORY

By J. FRENKEL. The first volume of the comprehensive treatise of which 'WAVE MECHANICS, ADVANCED 'GENERAL THEORY, is the second volume. 1932. Royal 8vo, pp. 286.

THE PRINCIPLES OF QUANTUM MECHANICS

By P. A. M. DIRAC. 1930. Royal 8vo, pp. 268.

'Dr. Dirac's book has been eagerly awaited. . . . A book taking stock of the whole position, by one of the leaders in the advance, raises hope that our difficulties may be settled. The book fills this need admirably. For the first time wave mechanics is presented in a really coherent form with something like a philosophy of the new methods to support it.'

SIR ARTHUR EDDINGTON in the *Cambridge Review.*

'There can be no doubt that his [Dr. Dirac's] work ranks as one of the high achievements of contemporary physics.'—*Nature.*

CONSTITUTION OF ATOMIC NUCLEI AND RADIOACTIVITY. By G. GAMOW. 1931. Royal 8vo, pp. 122.

'This is an important book which will be of great technical use to workers in the subject and in addition provide an informative and interesting account for more general readers.'—*Cambridge Review.*

'Il était donc utile de "faire le point" et c'est ce à quoi le livre de Gamow réussit parfaitement.'—*Journal de Physique.*

THE THEORY OF ELECTRIC AND MAGNETIC SUSCEPTIBILITIES. By J. H. VAN VLECK. 1932. Royal 8vo, pp. 396.

Professor Van Vleck, in this volume, has made a valuable contribution to the Literature of Physics and this book is well worthy of the attention of those interested in the subject.'—*Chemical News.*

THE THEORY OF ATOMIC COLLISIONS

By N. F. MOTT and H. S. W. MASSEY. 1933. Royal 8vo, pp. 298.

WAVE MECHANICS

ADVANCED GENERAL THEORY

BY

J. FRENKEL

PROFESSOR AT THE PHYSICO-TECHNICAL INSTITUTE, LENINGRAD

OXFORD
AT THE CLARENDON PRESS
1934

OXFORD UNIVERSITY PRESS
AMEN HOUSE, E.C. 4
LONDON EDINBURGH GLASGOW
LEIPZIG NEW YORK TORONTO
MELBOURNE CAPETOWN BOMBAY
CALCUTTA MADRAS SHANGHAI
HUMPHREY MILFORD
PUBLISHER TO THE
UNIVERSITY

PRINTED IN GREAT BRITAIN

PREFACE

THE present volume forming the second Part of my *Wave Mechanics* is devoted (as foreshadowed in the Preface to Part I) to the mathematical development of the general ideas underlying the new mechanics, connecting it with classical mechanics and constituting it a complete self-supporting theory. In building up the mathematical framework of this theory I have limited myself to what I consider its most essential elements, leaving aside a number of questions which have a methodological value only (such as the group theory) or which are met with in the solution of special problems.

It is my intention to consider some of these questions later on in connexion with the special problems which will be discussed in Part III ('Advanced Special Theory'); I have carefully avoided complicating the general scheme of the theory by such special questions—with a few exceptions inserted for illustration (the relativistic theory of the hydrogen-like atom, for example).

To make the general scheme more comprehensible I have not spared space, dealing with especially important general questions (such as the transformation and the perturbation theory, or the relativistic theory of the electron) at much greater length than would be necessary from the point of view of an adequate presentation to a sophisticated reader.

I must cordially thank the editors for their readiness to meet my demands on space, which have resulted in a book larger than was originally contemplated. I must also thank M. L. Urquhart and Miss B. Swirles for help in correcting the English and the proofs.

The present book, like Part I, is complete in itself, and can be read without acquaintance with Part I, provided the reader is familiar with some elementary account of wave mechanics, and is ready to explore its mathematical depths to obtain a profounder insight into the theory and to prepare himself for applying it to various special problems.

The earlier portions of this book were written in 1931 while I was in America; it was completed in Leningrad nearly two years later. Some of the shortcomings of the book are due to this interruption and the impossibility of revising it in 1933 from the very beginning.

A list of the more important references for each section is given at the end of the book; it is followed by a short index which should enable the reader to locate easily all the more important subjects treated.

LENINGRAD J. F.
Nov. 1933

CONTENTS

ADVANCED GENERAL THEORY

I

CLASSICAL MECHANICS AS THE LIMITING FORM OF WAVE MECHANICS

1. Motion in One Dimension; Partial Reflection and Uncertainty in the Sign of the Velocity

In the first part of this book we have given a general outline of the development and present state of wave mechanics, emphasizing the physical meaning of the new conceptions and avoiding, as far as possible, formal questions connected with the mathematical expression of these new conceptions. We have thus been led astray from the old conceptions based on classical corpuscular mechanics, deepening, as it were, the abyss separating the old from the new mechanics.

A systematic study of the formal questions referred to above reveals the wonderful fact that in spite of the fundamental physical difference between the new and the old mechanics, they are extremely similar from the mathematical point of view, i.e. from the point of view of the mathematical expression of the various physical quantities and the mathematical equations connecting them. This formal similarity forms a bridge over the abyss between the old and the new mechanics, enabling one to consider the latter as an extension or rather a refinement of the former and to establish a one-to-one correspondence between the old 'classical' and the new 'quantum' conceptions, quantities, and equations—a correspondence which often looks like an identity.

The existence of such a correspondence is a very instructive example of the fact—many times already illustrated by the development of physics—that a drastic revision of our physical conceptions can be associated with a simple improvement in the underlying mathematical scheme.

We shall start by considering the simplest case of the wave-mechanical equation, i.e. the equation describing the stationary motion of a particle in one dimension:

$$\frac{d^2\psi}{dx^2}+\frac{8\pi^2m}{h^2}(W-U)\psi=0, \tag{1}$$

the potential energy U being supposed to depend on x only (and not upon t, otherwise the total energy W would not be constant).

We should thus have the equation

$$\frac{d\phi}{dx} = \alpha, \tag{5}$$

which is inconsistent with (3 a) unless $d^2A/dx^2 = 0$. This condition, giving $A = ax+b$, is, however, in general inconsistent with the relation (4 a), i.e. $A^2\alpha = C$, unless $\alpha = \frac{C}{(ax+b)^2}$, which means a very special assumption for the potential-energy function U (the preceding relation is fulfilled in particular if $U = \text{const.}$, a being equal to zero in this case). We thus see that a one-sided wave propagation, corresponding to the motion of a particle in one definite direction, is in general impossible.

From the point of view of the wave conception this result is very easily explained. Thus every field of force, i.e. every inhomogeneity in the potential energy U or the parameter α, leads to a *partial reflection* of a wave impinging on it. If the inhomogeneity is due to a discontinuous jump of α, the reflection is produced at the point (or plane) of discontinuity. If α varies continuously, the reflection is produced gradually (the reflected waves giving rise to reflected waves of the second order travelling in the initial direction, and so on).

From the corpuscular point of view this means that a particle moving along the axis of x in a field of force parallel to x may have its velocity reversed at every instant, so that while the magnitude of the velocity is a given function of x, its *direction* or sign *remains uncertain*.

This uncertainty constitutes the fundamental difference between the new and the old mechanics. In the old mechanics, if the direction of the velocity is fixed at some initial instant, then it should remain the same so long as the kinetic energy $W-U$ remains positive ($\alpha^2 > 0$). Such a determinateness does not actually exist in the phenomena of motion. When these phenomena are described by wave mechanics, we find Nature in a position very similar to that of a theoretical physicist who, in performing complicated (and even simple!) calculations, often feels a strong uncertainty about the sign (+ or −) which must be assigned to the quantities under consideration.

This uncertainty of sign or of direction of velocity for a given magnitude of the latter and a given position can be regarded as an 'uncertainty principle' characteristic of wave mechanics and not related directly to the uncertainty principle of Heisenberg. The difference between them is that in the latter the localization of the particle is imagined to be effected by means of a 'wave packet' involving an uncertainty not so much

in the direction of the velocity as in its magnitude, whereas in the present case there is no need for constructing such a packet, the *fact asserted being not a definite position* of the particle, but the *connexion* between position, which may be arbitrary (that is, specifiable in terms of probability only) and the magnitude of the velocity. As we have just seen, the uncertainty in the direction of this velocity is connected with the possibility of both transmission and reflection of the particle in every region where it is acted on by some force. At the very beginning of this book we came upon this possibility when attempting to interpret, from the corpuscular point of view, the phenomena of partial reflection and partial transmission of light at the boundary between two homogeneous bodies. Later we studied it in more detail when investigating the motion of material particles in a field of force according to wave mechanics. We can sum up the results arrived at by saying that the indeterminateness which constitutes the characteristic distinction between wave mechanics and classical mechanics is due primarily to this ambiguity in the result produced by a force acting on the particle. Whereas in classical mechanics such a force must either accelerate or retard the particle, reversing the direction of its motion only when the increase of potential energy would exceed the total energy, in wave mechanics a force can reverse the direction of motion, leaving the magnitude of the velocity unchanged, even when this force is acting in the direction of the motion, i.e. even when, according to classical mechanics, the particle should be accelerated without change of direction.

So far as the relation between the wave-mechanical and the classical equations of motion is concerned, this uncertainty in the direction or in the 'sign' of the velocity, when its magnitude and the position of the particle are simultaneously fixed, is much more useful than Heisenberg's uncertainty principle (which is another aspect of the fundamental ambiguity inherent in wave mechanics). It leads us to expect that *the results predicted by wave mechanics will approach those predicted by classical mechanics as the reflection coefficient tends to zero*, i.e. when the ambiguity due to the possibility of reflection as well as transmission vanishes. In this case, transmission, i.e. motion in the same direction, is the only issue that comes into consideration.

It is easy to see that a decrease in the reflection coefficient is brought about by a *decrease in the wave-length*. When the wave-length becomes very small compared with the length over which the potential energy changes by an appreciable amount, the reflection produced by this

change of potential energy also becomes very small and vanishes in the limiting case $\lambda = 0$.

This result can be illustrated by the fact, pointed out in Part I, § 12, that cathode rays pass without appreciable reflection through an electric condenser whose thickness is very large compared with the wave-length, while they are appreciably reflected if this thickness is reduced to zero, the potential energy change remaining the same. In the latter case the reflection and transmission coefficients are given by the well-known formulae

$$R = \left(\frac{\alpha'-\alpha''}{\alpha'+\alpha''}\right)^2, \qquad D = 1-R = \frac{4\alpha'\alpha''}{(\alpha'+\alpha'')^2},$$

where α' and α'' are the values of the parameter α on both sides of the discontinuity. It may be recalled that this parameter is proportional to the momentum $g = mv$, i.e. to the velocity of the electron. When the velocity of the impinging electrons, that is α', increases, the jump ΔU of the potential energy remaining constant, α'' also increases, while the difference $\alpha'-\alpha''$ decreases. We have in fact, according to (3),

$$\Delta U = U''-U' = \frac{h^2}{8\pi^2 m}(\alpha'^2-\alpha''^2),$$

whence

$$\alpha'-\alpha'' = \frac{8\pi^2 m}{h^2}\frac{\Delta U}{\alpha'+\alpha''},$$

or approximately

$$\frac{\alpha'-\alpha''}{\alpha'+\alpha''} = \frac{8\pi^2 m}{h^2}\frac{\Delta U}{4\alpha^2} = \frac{\Delta U}{4(W-U)},$$

that is,

$$R \cong \frac{1}{16}\left|\frac{\Delta U}{W-U}\right|^2. \tag{5a}$$

Here $W-U$ is the average kinetic energy $\frac{1}{2}mv^2$ of the electron on both sides of the discontinuity, while ΔU is equal to the change of this kinetic energy, i.e. approximately $mv\,\Delta v$.

We thus get

$$R = \frac{1}{4}\left(\frac{\Delta v}{v}\right)^2 = \frac{1}{4}\left(\frac{\Delta\lambda}{\lambda}\right)^2, \tag{5b}$$

where

$$\lambda = \frac{h}{mv}.$$

Formula (5a) shows that the reflection coefficient tends to zero when the velocity of the electron is increased, i.e. when the wave-length λ tends to zero, the jump of potential energy ΔU remaining constant ($\Delta\lambda$ is an infinitely small quantity of a higher order than λ itself).

This result holds, of course, not only for electrons but also for any other particles: their behaviour conforms more and more to the funda-

mental principle of classical mechanics, the principle of determinism which can be stated in the form

$$R = 0, \qquad D = 1$$

as their velocity increases.

It should be noted that, for a given value of ΔU, the magnitude of the velocity for which R becomes inappreciable is the smaller *the larger the mass* m, since, according to (5 a), it is not the velocity itself but the kinetic energy $\frac{1}{2}mv^2$ whose ratio to ΔU determines R.

2. Comparison between the Schrödinger and the Classical Equation of Motion in One Dimension; Average Velocity and Current Density

Discontinuities in the potential-energy function $U(x)$ do not, of course, occur in Nature. When $U(x)$ is a continuous function of x, i.e. when the force has a finite value, it is possible to give another important and interesting formulation of the condition under which the fundamental ambiguity of wave mechanics disappears (i.e. the reflection coefficient vanishes), the wave mechanics thus reducing to classical mechanics. According to de Broglie's relation $\lambda = h/mv$, the wave-length of the waves associated with the motion of a particle is, other things being equal, the smaller, the smaller the value of the constant h. In reality, of course, the latter cannot be changed. If, however, it were not a universal constant, but could have any value whatsoever, then it would be possible to say that wave mechanics would reduce to classical mechanics in the limiting case $h = 0$; for this would mean that the wave-length would vanish for all values of the velocity. Consequently the relative change of the potential energy in a distance of the order of magnitude of the wave-length would also vanish, and with it the partial reflection which is the fundamental cause of the ambiguity characteristic of wave mechanics.

This result can be proved in a general way as follows:

Let us put $\alpha = 2\pi g/h$ in equation (3 a), where $g\,(= mv)$ is the magnitude of the momentum of the particle, and also

$$\phi = \frac{2\pi}{h} s. \tag{6}$$

Multiplying (3 a) by $(h/2\pi)^2$, we get

$$\left(\frac{h}{2\pi}\right)^2 \frac{d^2A}{dx^2} + \left[g^2 - \left(\frac{ds}{dx}\right)^2\right] A = 0, \tag{6 a}$$

where

$$g^2 = 2m(W - U). \tag{6 b}$$

It follows from this equation that in the limiting case $h = 0$ the function s remains finite and is determined by the differential equation

$$\left(\frac{ds}{dx}\right)^2 = 2m(W-U). \tag{7}$$

The momentum g can be determined by this function unambiguously, i.e. both with respect to magnitude *and sign*, by the equation

$$g = \frac{ds}{dx}, \tag{7a}$$

which is equivalent to equation (5), corresponding to the one-sided wave propagation, i.e. to the motion of a particle in a *definite* direction. This direction remains arbitrary, since (7) has two solutions, namely $ds/dx = +\sqrt{\{2m(W-U)\}}$ and $ds/dx = -\sqrt{\{2m(W-U)\}}$. But once it is chosen for some initial instant it will remain constant so long as s is a continuous function of x without maxima or minima, where, of course, g will change its sign after passing through the value $g = 0$. This change of sign through a continuous variation corresponds to total reflection and has nothing to do with the discontinuous reversal of the sign of g which is allowed by the exact theory embodied in the wave equation (1) (with $h > 0$) and which corresponds to *partial* reflection. The difference between the exact equation (1) and the approximate equation (7), so far as the ambiguity in the sign, i.e. in the direction of the velocity, is concerned, consists in the fact that the former, being a *linear* equation of the *second* order, admits *both signs simultaneously* (superposition of waves travelling in opposite directions), while the latter, being a *quadratic* equation of the *first* order, admits *either* one sign *or* the other. It should be remembered that the exact equation which is satisfied by the function s is much more complicated than (7). This exact equation can be obtained by eliminating A from equations (3a) and (3b) with $\phi = 2\pi s/h$.

It is often convenient to use, instead of the function defined in this way, another function S defined by the equation

$$\psi = e^{i2\pi S/h}, \tag{8}$$

or

$$S = \frac{h}{2\pi i}\log\psi. \tag{8a}$$

This S is connected with s (i.e. the 'phase' ϕ) and the 'amplitude' A by the relation

$$S = s + \frac{h}{2\pi i}\log A.$$

It is a complex quantity which represents both ϕ and A and is equivalent to ψ.

Substituting the expression (8) in Schrödinger's equation (1) and using the relations

$$\frac{d\psi}{dx} = i\frac{2\pi}{h}\frac{dS}{dx}e^{i2\pi S/h}; \qquad \frac{d^2\psi}{dx^2} = \left(\frac{i2\pi}{h}\right)^2\left(\frac{dS}{dx}\right)^2 e^{i2\pi S/h} + i\frac{2\pi}{h}\frac{d^2S}{dx^2}e^{i2\pi S/h},$$

we get

$$\frac{h}{2\pi i}\frac{d^2S}{dx^2} + \left(\frac{dS}{dx}\right)^2 = 2m(W-U). \tag{8 b}$$

If we put here $h = 0$, this equation reduces to (7), so that when $h = 0$ the two functions s and S become identical. We must now investigate the meaning of the approximate equation (7) which they both satisfy in this limiting case.

In a certain sense it merely expresses the law of the conservation of energy—since ds/dx is, by definition, the momentum g of the particle and $\frac{1}{2m}\left(\frac{ds}{dx}\right)^2$ is its kinetic energy.

The equation is unusual, however, in that the momentum of the particle, and consequently its velocity, is determined as a function of the coordinate x, whereas in the classical description of motion the velocity, as well as the coordinate itself, usually appear as functions of the time t. Such a description of motion is impossible in wave mechanics because of the uncertainty in the direction of the velocity. If it is true, however, that in the case $h = 0$ the wave-mechanical equation of motion (8 b) must reduce to the classical equation, then equation (7) must be equivalent to Newton's equation of motion

$$m\frac{d^2x}{dt^2} = -\frac{dU}{dx}, \tag{9}$$

defining x and $v = dx/dt$ as functions of the time. This equivalence is readily recognized as soon as we realize what is meant by defining the velocity (or momentum) of a particle as a function of its coordinate. Let us suppose that equation (9) has been integrated, and that x and v have been determined as functions of the time t. Then, eliminating the time t between them, we can express one of them, e.g. v, as a function $v(x)$ of the other. The acceleration d^2x/dt^2 can then be calculated by means of the formula

$$\frac{d^2x}{dt^2} = \frac{dv}{dt} = \frac{dv}{dx}\frac{dx}{dt} = \frac{dv}{dx}v = \frac{d}{dx}\left(\frac{v^2}{2}\right),$$

so that equation (9) can be written in the form

$$\frac{d}{dx}\frac{mv^2}{(2)} = -\frac{dU}{dx}$$

or

$$\frac{mv^2}{2} + U = \text{const.}$$

If $mv = g$ is replaced by ds/dx and the constant is denoted by W, we get equation (7).

We thus see that this equation expresses not only the law of conservation of energy, but at the same time the classical law of motion. It should be mentioned that both laws are equivalent to one another only in the special case which we are considering here of motion in *one dimension* (see below).

Another way of interpreting equation (7), or rather the fact implied in it that the velocity $v = \frac{1}{m}\frac{ds}{dx}$ of the particle is determined not as a function of the time but as a function of the coordinate x, is to replace the single particle under consideration by an infinite number of copies of this particle, filling space (or the line x) in a continuous way, so that at any instant t a copy is to be found situated at, or rather passing through, any point x. This method is similar to one used in hydrodynamics except that, in the hydrodynamical case, the copies of a particle are replaced by actual particles (supposed to be identical), moving under the combined influence of external forces and forces of mutual action (represented by the hydrostatic pressure). Provided we are not interested in the individuality of the particles, i.e. in the question *which* particle is to be found at a given point, the motion of the particles can be specified by defining the velocity of the particle passing through each fixed point as a function of the coordinates of this point and, in general, of the time. If the velocity does not depend upon the time (it should be remembered that the velocity we are speaking of refers not to a definite particle but to a definite point) the motion is called *stationary* or *steady*.

Thus the picture which can be associated with equation (7) is that of an assembly of copies of the particle under consideration, streaming steadily and filling space in a continuous way. If we select from this assembly a definite copy which at the time t was passing through the point x, then, knowing the dependence of the velocity v upon x, we can follow its motion and determine both the velocity and position of *this particular* copy as functions of the time. For instance, at the

moment $t+dt$ the copy in question will be situated at the point $x+v\,dt$, and will have the velocity $v(x+dx) = v(x+v\,dt) = v(x)+\frac{dv}{dx}v\,dt$, which means that its acceleration is equal to $v\,dv/dx$, as was obtained above.

We have thus shown that the wave-mechanical equation of motion actually reduces to the classical equation in the limiting case when the wave-length associated with the motion of a particle tends to zero, either owing to increase in velocity (which is a thing that can actually happen) or to decrease in the constant h (which is an artifice). The fundamental reason for this lies in the elimination of partial reflection, i.e. of a reversal in the direction of the velocity or, in other words, the elimination of the uncertainty in its sign.

Strictly speaking, however, this uncertainty cannot be eliminated. It is impossible to describe the motion of a particle in the classical way, i.e. as a determinate change of position and velocity with the time. The only way of describing it is to ascertain the *probability* of finding the particle at a given place and the probability that, being at this place, it is moving in the one or the other direction (the magnitude of the velocity being fixed). This intrusion of the probability conception into the description of the motion is necessary because of the ambiguity arising from the alternative: partial reflection or partial transmission. One could say that this ambiguity—wholly alien to classical mechanics —forms the gate through which the concept of probability penetrates into the realm of physics.

The probability of position is measured, as we know, by the product $\psi\psi^*$, so that $\psi(x)\psi^*(x)\,dx$ measures the probability that the particle is situated in the region between x and $x+dx$. Using the picture of an assembly of copies of the particle in question filling space (or the x-axis) in a continuous way, we can interpret $\psi\psi^*\,dx$ as the relative number of copies situated within the interval dx (this number is independent of the time so long as $\psi = \psi^0 e^{-i2\pi\nu t}$, corresponding to a motion with a definite total energy $W = h\nu$). If the integral $\int_{-\infty}^{+\infty}\psi\psi^*\,dx$ converges, ψ can be normalized in such a way that this integral is equal to 1, in agreement with the usual normalization of probability. Otherwise we need not worry about this normalization, since after all only relative values of $\psi\psi^*$ for different points come into account.

It should be noticed that in the classical description of the motion we can also use a continuous assembly of copies instead of an individual particle, as is actually done when the equation of motion is written in

the form (7) corresponding to the determination of the velocity as a function of the coordinate and not of the time. From the point of view of this description the difference between the old and the new theory can be summed up as follows. In the old theory it is always possible to 'individualize' a certain copy by following its motion, i.e. by determining its coordinate and velocity as definite functions of time, whereas in the new theory such 'individualization' is impossible, the direction of motion being uncertain. It thus becomes necessary to consider the assembly as a whole without attempting to disentangle it, i.e. to trace the motion of a particular copy in time. This being so, the density of the assembly, i.e. the relative number of copies per unit range, or, in other words, the probability of finding the particle represented by these copies in a given range, becomes the primary thing that can and must be determined—whereas in classical mechanics it remains irrelevant and therefore arbitrary. Of course the determination of $\psi\psi^*$ in wave mechanics is also connected with some arbitrariness, which can only be removed by specifying the boundary conditions or the conditions at infinity for the function ψ.

Knowing the function ψ, one can determine many other things besides the probability of position. Thus by means of it we can determine the probability of the two opposite directions of motion, that is, of the two opposite signs of the velocity, if the magnitude of the velocity is assumed to be fixed for a given position by the classical relation $v = \sqrt{\{2(W-U)/m\}}$ or by de Broglie's relation $v = h/(m\lambda)$. If p' is the probability of the positive direction and p'' that of the negative direction, then the average or probable value of the velocity at a given point is given by the formula

$$\bar{v} = (p'-p'')|v| \tag{10}$$

with the condition $p'+p'' = 1$.

This probable velocity, or the probabilities p, can be determined quite generally with the help of the relation (4), as soon as the physical meaning of this relation is recognized. We shall first see what the expression $A^2\, d\phi/dx$ means in the simple case of a wave travelling in one direction in a force-free space, that is, a wave representing the free motion of a particle in one direction. We have, in this case, according to (1 a), $\phi = \alpha x$ and consequently $A^2 \frac{d\phi}{dx} = A^2\alpha = |\psi|^2 \frac{2\pi g}{h} = \frac{2\pi m}{h}|\psi|^2 v$. If $|\psi|^2$ is interpreted as the (relative) density of the copies of the particle, then the product $|\psi|^2 v = j$ must obviously be defined as the

corresponding *current density*, i.e. the (relative) number of copies passing through the given point or plane $x = \text{const.}$ in the direction of v in unit time. If $|\psi|^2$ is interpreted as the probability density, then j can be defined as the probability current density, i.e. the probability that the particle will cross the plane $x = \text{const.}$ in unit time. The ratio $j/|\psi|^2$ is nothing else than the actual velocity of motion, which, in view of the fact that the direction of the motion is perfectly definite, coincides with the probable velocity $\bar{v}$ (p' or $p'' = 1$).

It is natural to extend the above interpretation of the expression $A^2\, d\phi/dx$ as a measure of the current density to any type of wave function ψ, for from this point of view the fact that $A^2\, d\phi/dx$ is constant (i.e. independent of x) simply means that the number of copies passing through different planes $x = x_1$ and $x = x_2$, say, is the same, just as if they were actual indestructible particles. The law expressed by the relation (4) would thus be the law of conservation of the number of copies or of the *conservation of probability* (see below). If this interpretation is correct, then it must obviously be possible to write j in the form

$$j = \psi\psi^*\bar{v}, \tag{10a}$$

where $\bar{v}$ denotes the probable velocity of the copies at the point in question. Now this is actually the case if j is defined as $\frac{h}{2\pi m} A^2 \frac{d\phi}{dx}$ (the coefficient $h/2\pi m$ is the same as in the special case considered above), which gives the following expression for the probable velocity

$$\bar{v} = \frac{h}{2\pi m}\frac{d\phi}{dx}. \tag{10b}$$

The 'phase' ϕ can be expressed in terms of the function $\psi = Ae^{i\phi}$ and its conjugate complex $\psi^* = Ae^{-i\phi}$ by means of the formula

$$\phi = \frac{1}{2i}\log(\psi/\psi^*);$$

whence it follows that

$$\bar{v} = \frac{h}{4\pi i m}\left(\frac{1}{\psi}\frac{d\psi}{dx} - \frac{1}{\psi^*}\frac{d\psi^*}{dx}\right) = \frac{h}{2\pi m}\mathrm{R}\left(\frac{1}{i}\frac{d}{dx}\log\psi\right) \tag{10c}$$

or, according to (8a), $\bar{v} = \frac{1}{m}\mathrm{R}\left(\frac{dS}{dx}\right)$,

$\mathrm{R}(f)$ denoting the real part of f. In the classical theory this equation reduces to $\bar{v} = v$, in accordance with the fact that the motion proceeds in a perfectly definite direction, the probabilities p' and p'' being equal respectively to 1 and 0. In the wave-mechanical theory $|\bar{v}|$ is, in general,

different from $|v|$, the values of the probabilities p' and p'' being different from both 1 and 0. They can be determined from v and $\bar{v}$ by means of the formula

$$p = \frac{1}{2}\left(1 \pm \left|\frac{\bar{v}}{v}\right|\right).$$

Substituting (10 c) in (10 a), we get the following expression for the current density:

$$j = \frac{h}{4\pi i m}\left(\psi^* \frac{d\psi}{dx} - \psi \frac{d\psi^*}{dx}\right) = \frac{h}{2\pi m} \mathrm{R}\left(\frac{1}{i}\psi^* \frac{d\psi}{dx}\right). \tag{11}$$

We shall now check these results by applying them to two simple cases.

We shall put first

$$\psi = A' e^{i\alpha x} + A'' e^{-i\alpha x},$$

which corresponds to the free motion of a particle along the x-axis in an unspecified direction.

Assuming the coefficients A for the sake of simplicity to be real (this condition does not involve any loss of generality, for it can always be satisfied by a suitable choice of the origin $x = 0$), we have

$$\psi^* = A' e^{-i\alpha x} + A'' e^{i\alpha x},$$

whence

$$\begin{aligned} \frac{1}{i}\psi^* \frac{d\psi}{dx} &= \alpha(A'^2 - A''^2) + \alpha(A'A'' e^{i2\alpha x} - A'A'' e^{-i2\alpha x}) \\ &= \alpha(A'^2 - A''^2) + i2\alpha A'A'' \sin 2\alpha x, \end{aligned}$$

so that j reduces to the constant value,

$$j = \frac{h\alpha}{2\pi m}(A'^2 - A''^2)$$

or

$$j = |v|(A'^2 - A''^2). \tag{11 a}$$

Unlike j, the probable velocity

$$\bar{v} = \frac{j}{\psi\psi^*} = |v| \frac{A'^2 - A''^2}{A'^2 + A''^2 + 2A'A'' \cos 2\alpha x}$$

is a function of x, varying periodically between the values

$$\bar{v}_{\max} = |v| \frac{A' + A''}{A' - A''}$$

and

$$\bar{v}_{\min} = |v| \frac{A' - A''}{A' + A''}.$$

The fact that the maximum value of the probable velocity $\bar{v}$ turns out to be *larger* than the magnitude of the classical velocity $|v|$ invalidates the idea considered above of taking the latter over into the wave-mechanical theory as the magnitude of the 'actual' velocity. With

$|\bar{v}/v| > 1$ formula (10) leads to values of the probabilities p which are devoid of physical meaning, one of them being larger than 1 and the other smaller than 0. Although the classical velocity can be determined wave-mechanically from the wave-length λ (by means of the formula $|v| = h/(m\lambda)$), yet it is the probable velocity $\bar{v}$ only which has a direct physical significance.

This is also clearly seen if we take as a second example the case

$$\psi = A'e^{+\beta x} + A''e^{-\beta x}$$

corresponding to a region of *total reflection* where the kinetic energy is negative and the velocity v is imaginary. We have in this case $\psi^* = \psi$, $j = 0$, and $\bar{v} = 0$, as might be expected.

3. Generalization for Non-stationary Motion in Three Dimensions; The Hamilton-Jacobi Equation

We shall now generalize the results of the preceding section to the motion of a particle in three dimensions under the action of forces derived from a potential-energy function U which may depend not only upon the coordinates x, y, z, but also upon the time t.

The wave-mechanical description of such a motion is given by the generalized equation of Schrödinger

$$\nabla^2\psi - \frac{8\pi^2 m}{h^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} + U\right)\psi = 0. \tag{12}$$

Our main object will be to trace the relation of this equation to the corresponding classical equations of motion,

$$m\frac{d^2x}{dt^2} = -\frac{\partial U}{\partial x}, \qquad m\frac{d^2y}{dt^2} = -\frac{\partial U}{\partial y}, \qquad m\frac{d^2z}{dt^2} = -\frac{\partial U}{\partial z}. \tag{12a}$$

The general character of this relation can be described in a way similar to that used for the one-dimensional motion discussed above. The fundamental characteristics of the wave-mechanical theory can thus be partially reduced, as before, to the ambiguity arising from the phenomenon of partial reflection and partial transmission—a phenomenon which implies a sudden change in the direction of the velocity, its magnitude being assumed to be the same function of the coordinates as in the classical theory.

The uncertainty in the direction of the velocity, which in the case of one-dimensional motion was equivalent to an ambiguity of sign, is now—in the case of motion in space—of a still more distressing character. However, we may still expect this uncertainty, as well as partial reflection, to vanish in the limiting case of motion corresponding

to infinitely short wave-lengths (which can be realized by an increase of velocity or of mass, or by a fictitious decrease of the constant h). Thus in this limiting case equation (12) must become equivalent to equations (12 a) in the sense of admitting *particular* solutions corresponding to a perfectly definite type of classical motion.

To demonstrate this equivalence we shall replace the particle under consideration by an assembly of copies distributed and moving in space like the particles of some continuous fluid (without interaction of course!). The velocity vector $\mathbf{v}$ of each copy can then be defined—according to the classical theory—as a function of the coordinates x, y, z of the (fixed) point through which this copy is passing, and of the time—the motion being not necessarily a steady one. It should be noticed that the partial derivative $\partial\mathbf{v}/\partial t$ of $\mathbf{v}$ with regard to the time does not define the acceleration of a given copy, for it refers to different copies passing through the same point at different instants of time t and $t+dt$. This acceleration can be defined by the total derivative $d\mathbf{v}/dt$, its x-component being thus given by

$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x}\frac{dx}{dt} + \frac{\partial v_x}{\partial y}\frac{dy}{dt} + \frac{\partial v_x}{\partial z}\frac{dz}{dt}$$

or

$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + v_x\frac{\partial v_x}{\partial x} + v_y\frac{\partial v_x}{\partial y} + v_z\frac{\partial v_x}{\partial z}. \tag{13}$$

We shall now assume the motion of the fluid formed by our assembly of copies to be *irrotational*, which means that the velocity vector can be represented as the gradient of a scalar function, the so-called 'velocity potential'. We shall denote this function by s/m and put accordingly

$$m\mathbf{v} = \nabla s, \tag{13 a}$$

that is

$$v_x = \frac{1}{m}\frac{\partial s}{\partial x}, \qquad v_y = \frac{1}{m}\frac{\partial s}{\partial y}, \qquad v_z = \frac{1}{m}\frac{\partial s}{\partial z}.$$

We make this assumption (which is by no means necessary) not only because we desire to simplify the formulation of the classical theory as applied to the copy assembly, but also because we wish to establish the connexion between this theory and the wave-mechanical theory. We have in fact, for a wave propagated in one definite direction, a relation exactly similar to (13 a) between the phase ϕ and the vector $\boldsymbol{\alpha}$ whose direction is the direction of propagation and whose length is $2\pi/\lambda$, where λ is the value of the wave-length at the corresponding point:

$$\boldsymbol{\alpha} = \nabla\phi. \tag{14}$$

If we put
$$\alpha = 2\pi\frac{m\mathbf{v}}{h}, \tag{14a}$$
according to de Broglie's relation, we get
$$\phi = \frac{2\pi}{h}s \tag{14b}$$
as before [cf. (6), § 1]. Thus, by assuming irrotational motion of the assembly of copies, it becomes possible to establish a connexion between the motion of a particle and the propagation of waves in the limiting case of infinitely short waves, i.e. when partial reflection is excluded and the motion of every copy of the particle proceeds along a perfectly definite path; this path can be considered as the 'ray' passing through the point at which the copy in question was initially situated. If partial reflection does take place the idea of rays loses all meaning, each ray branching into two at every point. Only by neglecting reflection can one speak of rays as lines along which the waves, i.e. the surfaces of constant phase, are propagated.

Returning to the expression (13) for the x-component of the acceleration of the copy passing through the point x, y, z, at the instant t we can, because of (13a), rewrite it in the form
$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + v_x\frac{\partial v_x}{\partial x} + v_y\frac{\partial v_y}{\partial x} + v_z\frac{\partial v_z}{\partial x},$$
since $\frac{\partial v_x}{\partial y} = \frac{1}{m}\frac{\partial^2 s}{\partial x \partial y} = \frac{\partial v_y}{\partial x}$, etc. Therefore
$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + \frac{\partial}{\partial x}\left(\frac{v^2}{2}\right)$$
or
$$\frac{dv_x}{dt} = \frac{1}{m}\frac{\partial}{\partial x}\left[\frac{\partial s}{\partial t} + \frac{1}{2m}(\nabla s)^2\right].$$
The equation $m\frac{dv_x}{dt} = -\frac{\partial U}{\partial x}$, which is the first of the equations (12a), is thus equivalent to
$$\frac{\partial}{\partial x}\left[\frac{\partial s}{\partial t} + \frac{1}{2m}(\nabla s)^2 + U\right] = 0.$$
Similar results are obtained for the second and the third equations, and so all three of them can be replaced by the single equation
$$\frac{\partial s}{\partial t} + \frac{1}{2m}(\text{grad}\, s)^2 + U = F(t),$$
where $F(t)$ is an arbitrary function of the time alone. This function, without loss of generality, can be put equal to zero, for it corresponds

to an additive term $\int F(t)\,dt$ in s which is irrelevant for the determination of the velocity according to (13 a). The function s can thus be defined by the equation

$$\frac{\partial s}{\partial t}+\frac{1}{2m}\left[\left(\frac{\partial s}{\partial x}\right)^2+\left(\frac{\partial s}{\partial y}\right)^2+\left(\frac{\partial s}{\partial z}\right)^2\right]+U=0. \tag{15}$$

This equation was established by Hamilton and Jacobi and bears their name. In the special case when U does not depend upon the time explicitly (constant field of force), the function s—usually called the (mechanical) 'action'—reduces to

$$s=s_0(x,y,z)-Wt, \tag{15 a}$$

where s_0 is determined by the equation

$$\frac{1}{2m}\left[\left(\frac{\partial s_0}{\partial x}\right)^2+\left(\frac{\partial s_0}{\partial y}\right)^2+\left(\frac{\partial s_0}{\partial z}\right)^2\right]+U=W. \tag{15 b}$$

Here W is a constant which can obviously be defined as the *energy*. Thus, in a sense, equation (15 b), in conjunction with the relation (13 a), expresses the law of the conservation of energy. However, as we have just seen, it expresses much more than that,† since, in conjunction with (13 a), it is equivalent to the three classical equations of motion (12 a) for the special case of an invariable field of force and of a fixed value of the total energy. The equations (12 a) and (15 b)—or more generally (15)—are formally different because the former refer to an individual particle, while the latter refer to a continuous assembly of copies of this particle. If we select a definite copy and follow its motion we come back to equations (12 a).

It can now easily be shown that in the limiting case of infinitely small wave-length the wave equation (12) admits particular solutions of the form $\psi=Ae^{i\phi}$, representing a one-sided propagation of waves which can be associated, by means of the relations (14), (14 a), and (14 b), with the motion of the particle in question according to the classical theory, the different 'rays' coinciding with the paths of the different copies of this particle.

Putting $\psi=Ae^{i\phi}$, we get in the same way as in § 1

$$\frac{\partial^2\psi}{\partial x^2}=\frac{\partial^2 A}{\partial x^2}e^{i\phi}+2ie^{i\phi}\frac{\partial A}{\partial x}\frac{\partial\phi}{\partial x}+iAe^{i\phi}\frac{\partial^2\phi}{\partial x^2}-Ae^{i\phi}\left(\frac{\partial\phi}{\partial x}\right)^2,$$

whence

$$\nabla^2\psi=\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}=e^{i\phi}[\nabla^2 A-A(\nabla\phi)^2+i(2\nabla A\cdot\nabla\phi+A\nabla^2\phi)].$$

† Except in the one-dimensional case.

We have further

$$\frac{\partial \psi}{\partial t} = \frac{\partial A}{\partial t} e^{i\phi} + A e^{i\phi} \frac{\partial \phi}{\partial t}.$$

Substituting these expressions in equation (12), cancelling the common factor $e^{i\phi}$, and separating the real and imaginary parts, we obtain the two equations:

$$\nabla^2 A + \left[\frac{8\pi^2 m}{h^2}\left(-\frac{h}{2\pi}\frac{\partial \phi}{\partial t} - U\right) - (\nabla\phi)^2\right] A = 0$$

and

$$\frac{4\pi m}{h} \frac{\partial A}{\partial t} + 2\nabla A \cdot \nabla\phi + A\nabla^2\phi = 0.$$

If ϕ is replaced by $\frac{2\pi}{h} s$, these equations become

$$-\frac{h^2}{8\pi^2 m}\nabla^2 A + \frac{\partial s}{\partial t} + \frac{1}{2m}(\nabla s)^2 + U = 0, \tag{16}$$

$$2m\frac{\partial A}{\partial t} + 2\nabla A \cdot \nabla s + A\nabla^2 s = 0. \tag{16a}$$

Putting $h = 0$ we see that the first of these equations reduces to the Hamilton-Jacobi equation (15). The same result is obtained if $\nabla^2 A = 0$, which must obviously express the general condition for one-sided propagation of waves of finite length. In both cases the wave-mechanical theory becomes completely equivalent to the classical theory. Both cases are, of course, fictitious, h being a constant and the equation $\nabla^2 A = 0$ being satisfied only under very special conditions—in particular for force-free motion. The equation (16) can, however, reduce *approximately* to (15) in the case of a nearly one-sided wave propagation with a very weak partial reflection—so weak that the reflected (or scattered) waves can be neglected. This condition is more nearly approached the larger the mass m of the particle for a given velocity or the larger the velocity for a given mass, i.e. the smaller the wave-length, if we are treating motion corresponding to a constant value of the energy W. In the latter case the wave-length becomes a definite function of the coordinates. In the general case the idea of wave-length has no precise meaning and can be introduced only by representing the wave function ψ as a superposition of waves with different frequencies, corresponding to motions with different energies.

If U does not contain the time explicitly, equations (16) and (16a) admit particular solutions of the type $s = s_0(x, y, z) - Wt$ and

$A = A(x, y, z)$, i.e. $\partial s/\partial t = -W$ and $\partial A/\partial t = 0$. They therefore reduce to

$$-\frac{h^2}{8\pi^2 m}\nabla^2 A + \frac{1}{2m}(\nabla s_0)^2 + U = W \tag{17}$$

and

$$2\nabla A \cdot \nabla s_0 + A\nabla^2 s_0 = 0. \tag{17 a}$$

In the limiting case $h = 0$ the first of these becomes equivalent to the classical equation (15 b).

This equivalence, as well as the approximate equivalence which can be obtained in the case of large values of W or m, must not be misunderstood. It refers to *particular* solutions of equations (17) and (17 a), or of the corresponding Schrödinger equation

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(W - U)\psi = 0 \tag{17 b}$$

with

$$\psi = Ae^{i2\pi(s_0 - Wt)/h} = \psi^0(x, y, z)e^{-i2\pi Wt/h} \tag{17 c}$$

that is, to *solutions which represent—approximately—waves travelling in a definite direction* (the direction may, of course, vary from point to point, being defined by the direction of the 'rays' passing through these points). Now the general solution of (17 b) in the case of short waves can be represented as a superposition of a number of such particular solutions corresponding to waves travelling in different directions, under the limitations imposed by boundary conditions (in the case of long waves this is possible for force-free motion only). The classical equation (15 b), on the other hand, does not admit of such superposition for the function ψ defined as $Ae^{i2\pi s/h}$. This can clearly be seen in the simple case of one-dimensional motion where A is connected with s by the relation $A^2\frac{ds}{dx} = C$ [cf. (4), § 1], so that $\psi = \frac{\sqrt{C}}{\sqrt{(ds/dx)}}e^{i2\pi s/h}$. The physical reason for this is that 'superposition' of two different types of motion would mean, according to classical mechanics, their 'simultaneous realization'—an obviously impossible thing if they are alternative. In wave mechanics, on the contrary, it is just this alternative character which is expressed by superposition, the latter corresponding to the addition law of the classical probability theory. Similar results apply to the general equations (12) and (15), the former allowing the superposition of processes with different energies if U does not depend upon the time—while the latter reduces in this case to equation (15 b) corresponding to one definite value of the energy W.

The non-validity of the superposition principle in classical mechanics can easily be demonstrated with the help of the function $S = \frac{h}{2\pi i}\log\psi$

introduced in § 2 [eq. (8)]. This function satisfies the differential equation

$$\frac{h}{4\pi i m}\nabla^2 S+\frac{\partial S}{\partial t}-\frac{1}{2m}(\nabla S)^2+U=0 \tag{18}$$

which is obtained from Schrödinger's equation (12) by the substitution $\psi = e^{i2\pi S/h}$ and which reduces to the Hamilton-Jacobi equation (15) if h is put equal to zero. The function S thus coincides in this case with the function s, which means that the amplitude A can be considered as practically constant.

Now if in the Hamilton-Jacobi equation (15) we put $s = S = \frac{h}{2\pi i}\log\psi$, we get the following 'approximate' equation for ψ:

$$\frac{h}{2\pi i}\psi\frac{\partial\psi}{\partial t}-\frac{h^2}{8\pi^2 m}(\nabla\psi)^2+U\psi^2=0$$

or

$$(\nabla\psi)^2-\frac{8\pi^2 m}{h^2}\left(\frac{h}{4\pi i}\frac{\partial}{\partial t}+U\right)\psi^2=0, \tag{18a}$$

which is *quadratic* and of the first order (like the equation for S) instead of being *linear* and of the second order like the exact equation of Schrödinger. If ψ_1 and ψ_2 are two particular solutions of (18a), the function $\psi = \psi_1+\psi_2$ will not in general represent a solution of this equation.

Returning to the representation of the exact wave function in the form $Ae^{i\phi} = Ae^{i2\pi s/h}$, and considering equation (16a) connecting A and s, which has been disregarded hitherto, we see that this equation can be simplified if multiplied by A. We have in fact $2A\frac{\partial A}{\partial t} = \frac{\partial A^2}{\partial t}$ and

$$2A\nabla A\cdot\nabla s+A^2\nabla^2 s = \nabla(A^2)\cdot\nabla s+A^2\nabla^2 s = \operatorname{div}(A^2\nabla s);$$

so that

$$\frac{\partial(A^2)}{\partial t}+\operatorname{div}\left(A^2\nabla\frac{s}{m}\right)=0. \tag{19}$$

This equation is of the same form as the equation of continuity, i.e. the equation of the conservation of mass in hydrodynamics or of the conservation of electricity in electrodynamics,

$$\frac{\partial\rho}{\partial t}+\operatorname{div}\mathbf{j}=0,$$

where ρ is the density of mass or electrical charge and $\mathbf{j}$ the corresponding current density. In the present case we can interpret the quantity

$$A^2 = \psi\psi^* = \rho$$

as the density of the copy assembly (i.e. the relative number of copies

of the given particle in unit volume) or the density of probability. If, further, we define the corresponding *current density* by the formula

$$\mathbf{j} = \frac{1}{m} A^2 \nabla s, \tag{19a}$$

then equation (19) will express the law of the conservation of the copies or of the probability. In the classical theory the vector $\nabla s/m$ reduces to the actual velocity $\mathbf{v}$ of the particle (or more exactly of its copies at the given point), so that $\mathbf{j}$ assumes the usual form of the product of ρ with $\mathbf{v}$. In the exact wave-mechanical theory it can also be written in the form

$$\mathbf{j} = \rho \mathbf{v}$$

where the vector

$$\mathbf{v} = \frac{1}{m} \nabla s \tag{19b}$$

must obviously be interpreted as the *probable* velocity. The classical velocity can be computed as usual by means of the formula

$$v = \sqrt{\left\{\frac{2}{m}(W-U)\right\}},$$

its direction being, however, uncertain. According to the definition of A and s, we have $\psi = Ae^{i2\pi s/h}$, $\psi^* = Ae^{-i2\pi s/h}$, whence

$$s = \frac{h}{4\pi i} \log \frac{\psi}{\psi^*},$$

and consequently

$$\mathbf{j} = \frac{h}{4\pi i m}(\psi^* \nabla \psi - \psi \nabla \psi^*) = \frac{h}{2\pi m} \mathrm{R}\left(\frac{1}{i} \psi^* \nabla \psi\right). \tag{20}$$

Introducing the function $S = \frac{h}{2\pi i} \log \psi$, we get $\nabla S = \frac{h}{2\pi i} \frac{1}{\psi} \nabla \psi$ and $\frac{1}{i} \psi^* \nabla \psi = \frac{2\pi}{h} \psi \psi^* \nabla S$, so that

$$\mathbf{j} = \frac{1}{m} \psi \psi^* \mathrm{R}(\nabla S) \tag{20a}$$

and

$$\mathbf{v} = \frac{1}{m} \mathrm{R}(\nabla S). \tag{20b}$$

Comparing this with (19b), we see that the function s is equal to the real part of S, in accordance with the relation $S = s + \frac{h}{2\pi i} \log A$ which results from comparing the two expressions $e^{i2\pi S/h}$ and $Ae^{i2\pi s/h}$ for ψ. The probable velocity (20b) could be represented in the form

$$\mathbf{v} = |v| \int \mathbf{n} p(\mathbf{n})\, d\omega,$$

where $\mathbf{n}$ is the unit vector which defines the direction of the classical velocity and $p(\mathbf{n})\,d\omega$ is the probability that this unit vector lies in the infinitely small solid angle $d\omega$. An unambiguous determination of this probability appears, however, to be impossible, except for one-dimensional motion considered in the preceding section. This is quite natural if we remember that the notion of classical velocity, as measured by the time derivative of the coordinates, cannot be taken over into wave mechanics.

It should be mentioned in conclusion that the relation between wave mechanics and classical mechanics is usually compared with the relation between wave optics and the so-called geometrical optics, the latter being defined as the limiting case of wave optics for very small wave-lengths. This statement would, however, be misleading unless we add to it that in geometrical optics partial reflection of light (which actually decreases with decrease of wave-length) should be wholly left out of account—even in its simplest form on the boundary surface between two homogeneous media. In this case—and only in this case—is it possible to introduce the idea of rays as lines along which the propagation of light takes place (this is why geometrical optics is often called 'ray optics' in contradistinction to wave optics, where the idea of 'rays' has in general no meaning). It was the merit of Hamilton to show, one hundred years ago, that in this limiting case the wave conception of light can be replaced by the corpuscular conception, and that the rays can be described as the paths of light particles moving, according to Newton's classical law, in a certain field of force. The potential energy of this field of force U is determined by the refractive index μ according to the relation

$$\mu^2 = \gamma^2(W-U),$$

where γ is a constant depending upon the definition of the mass of a light particle.† But perhaps the main merit of Hamilton's work was that he applied the same considerations to the motion of particles of ordinary matter, thus *for the first time associating such motion with the propagation of* (infinitely short) *waves* and describing it by equation (15). This association of particles with waves, which in Hamilton's theory was achieved by interpreting the 'mechanical action' s as a measure of the phase function ϕ, was, however, completely forgotten for a hundred years, until de Broglie rediscovered it in the way described

† This relation is obtained in the simplest way by comparing de Broglie's formula for the wave-length $1/\lambda = \sqrt{\{2m(W-U)\}}/h$ with the formula $\lambda_0/\lambda = \mu$, which can be considered as the definition of the refractive index, λ_0 being the value of λ *in vacuo*, i.e. for a place where $\mu = 1$.

in Part I, and Schrödinger introduced his wave equation, whose relation to the Hamilton-Jacobi equation has been discussed above.

This mutual reaction of optics and mechanics must not be misinterpreted as an indication of a true analogy between them—in the sense of a wave-corpuscular duality of light. We must not be led by it to infer the real existence of photons, moving in material bodies according to the laws of wave mechanics. For we could replace optics by acoustics, i.e. light vibrations by mechanical vibrations propagated in the form of waves in elastic media according to an equation of exactly the same kind as the differential equation for the light waves. In the limiting case of infinitely short acoustical waves we could therefore obtain exactly the same results as in optics, i.e. a kind of 'ray acoustics' instead of a 'wave acoustics'. This would enable one to formulate a corpuscular theory of sound and describe the propagation of sound as the motion, according to wave mechanics, of certain particles—e.g. 'phonons'. I do not think, however, that anybody would believe in the reality of such 'phonons'. This does not mean, of course, that the photons are equally unreal, for the analogy between acoustics and optics is just as superficial as that between optics and mechanics (or acoustics and the mechanics of single particles).—I am inclined, however, to think that photons have no more reality than 'phonons', and that they are created by a 'reflection', as it were, of the wave-corpuscular duality of matter in the phenomena of light (cf. Part I).

4. Comparison of the Approximate Solutions of Schrödinger's Equation; Comparison of Classical and Wave-mechanical Average Values

Although in the case $h = 0$ the functions s and S satisfy the same equation—namely, that of Hamilton and Jacobi—yet the approximate expressions for ψ obtained therefrom, according to the formulae $\psi = Ae^{i2\pi s/h}$ and $\psi = e^{i2\pi S/h}$, turn out to be somewhat different, for the 'amplitude' A obtained by means of equation (16a) is in general a certain function of the coordinates (and the time), varying very slowly compared with the 'phase factor' $2\pi s/h$.

The discrepancy between the two approximate solutions is due to the fact that the error introduced by putting $h = 0$ is larger in the case of equation (18), which contains h in the first power, than in the case of equations (16) and (16a), where h appears in the second power. In the latter case we thus drop a small term of the second order, while in the former case we drop a much larger term of the first order.

In order to remove this discrepancy we must put

$$S = S^0 + \frac{h}{2\pi i} S', \tag{21}$$

and after substituting this expression in equation (18) drop terms which are quadratic in h but keep those which are linear in h (S^0 and S' being independent of h and therefore of the same order of magnitude). We thus get the approximate equation

$$\frac{h}{4\pi i m}\nabla^2 S^0 + \frac{\partial S^0}{\partial t} + \frac{h}{2\pi i}\frac{\partial S'}{\partial t} + \frac{1}{2m}(\nabla S^0)^2 + \frac{h}{2\pi m i}\nabla S^0 \cdot \nabla S' + U = 0. \tag{21a}$$

Here S^0 must be regarded as the zero approximation, corresponding to $h = 0$, i.e. as the solution of the Hamilton-Jacobi equation

$$\frac{\partial S^0}{\partial t} + \frac{1}{2m}(\nabla S^0)^2 + U = 0.$$

It can obviously be identified with the (approximate) function s.

The function S' must therefore satisfy the equation

$$\frac{1}{2m}\nabla^2 S^0 + \frac{\partial S'}{\partial t} + \frac{1}{m}\nabla S^0 \cdot \nabla S' = 0, \tag{21b}$$

whence it follows that S' is a real quantity. Now according to (21) we have $\psi = e^{i2\pi S/h} = e^{S'} e^{i2\pi S^0/h}$, so that, since $S^0 = s$, $e^{S'}$ must be equal to A. Substituting in (21 b)

$$S' = \log A, \tag{21c}$$

we do indeed get equation (16 a). It may seem that by developing the function S in a series of powers of the parameter $h/(2\pi i)$

$$S = S^0 + \frac{h}{2\pi i} S' + \left(\frac{h}{2\pi i}\right)^2 S'' + \ldots$$

and solving the equation (18) by successive approximations, one can obtain as good an approximation for S as may be desired. This assumption is, however, incorrect, for it can be shown that the preceding series is divergent or rather semi-convergent, which explains why one gets a closer approximation by keeping the first-order term, as has been done above. In fact the general solution of a differential equation of the second order cannot be approximated to by starting with the solution of the equation of the first order obtained by dropping the second-order terms, however small the parameter by which they are multiplied may be, just as a quadratic equation cannot be approximated to by the linear one obtained by dropping the quadratic term. If, however, the latter is multiplied by a small parameter, then

one of the two solutions of the quadratic equation can be approximated to by the solution of the linear one. A similar relationship exists between the function $\psi = e^{i2\pi S^0/h+S'}$ and one of the particular solutions of Schrödinger's equation, representing approximately waves travelling in one direction. It should be mentioned that this direction need not remain constant; it can be changed by *total* reflection, which, in contradistinction to partial reflection, is a phenomenon perfectly compatible with classical mechanics since it does not involve any ambiguity and therefore does not challenge a deterministic description of the motion. The difference between classical mechanics and wave mechanics in the approximate form given above, in so far as total reflection is concerned, consists only in the fact that, according to the latter, the particle can penetrate into those regions of the field of force where its 'classical' velocity becomes imaginary.

According to the relation $\mathbf{v} = \nabla s/m = \nabla S^0/m$, it should follow that the functions s and S^0 must also become imaginary. So far as S^0 is concerned this is perfectly true. The function s, however, according to its definition, must remain real. It will therefore be different from S^0 for those regions where v is imaginary and will satisfy an equation different from that of Hamilton and Jacobi. We must remember that equations (16) and (16a) were obtained on the assumption that both s and A were real. The assumption that s satisfies approximately the Hamilton-Jacobi equation, even when the latter gives imaginary values for it, would thus imply a contradiction.

This means that, in the case under consideration, $\nabla^2 A$ must be very large and of the order of magnitude of $1/h^2$, so that the first term in equation (16) or (17), which when omitted reduces (16) or (17) to the Hamilton-Jacobi equation, cannot be dropped. We shall not consider the approximate solution of equations (16) or (16a) [or (17) and (17a)] for this case. It is simpler to use instead the alternative representation of ψ by means of the function $S = S^0 + S'h/(2\pi i)$ since we do not have to worry about the reality of S^0. An imaginary value of S^0 leads, according to (21b), to an imaginary value of S'. The role of the functions S^0 and S' as determining the phase and the amplitude respectively will thus be reversed for classically forbidden regions, so that, using the expression $Ae^{i2\pi s/h}$ for ψ, we can put

$$A = e^{i2\pi S^0/h} = e^{\pm 2\pi |S^0|/h} \tag{22}$$

and

$$s = \frac{h}{2\pi i} S' = \pm \frac{h}{2\pi} |S'|. \tag{22a}$$

The sign (+ or −) is determined by the condition that A (i.e. ψ) must decrease with increased penetration into the forbidden region. It can easily be proved directly that the expressions (22) and (22 a) constitute an approximate solution of the equations (16) and (16 a) for the case in question if the functions S^0 and S' are determined respectively by the Hamilton-Jacobi equation and by equation (21 b).

Returning to the case when S^0 is real (and equal to s), corresponding to the motion in the classically allowed region of the field of force, let us examine the approximate values which are obtained for the amplitude $A = e^{S'}$.

We shall first consider the simplest case of a one-dimensional motion with constant energy. We have in this case, according to (4),

$$A^2 \frac{dx}{ds} = \text{const.},$$

that is, since

$$\frac{ds}{dx} = v,$$

$$A^2 = \frac{C^2}{|v|}, \tag{23}$$

where C^2 denotes a positive number. We thus get approximately

$$\psi = \frac{C}{\sqrt{|v|}} e^{i\frac{2\pi}{h}[s_0(x) - Wt]}, \tag{23 a}$$

$s_0(x)$ being a solution of the equation

$$\frac{1}{2m}\left(\frac{ds_0}{dx}\right)^2 + U = W.$$

Formula (23) has a very simple physical meaning. It shows that the probability of finding the particle within a certain region between x and $x+dx$ is inversely proportional to its velocity in this region. This is just what we should expect if this probability were defined as proportional to the time $dt = dx/v$ which the particle spends in the region in question. We thus see that the interpretation of the quantity $\psi\psi^* dx = A^2 dx$ as the relative probability of finding the particle in the region dx is in agreement, so far as the approximate expression for ψ is used, with the classical definition of probability in terms of *duration*.

If $f(x)$ is some quantity depending upon the position of the particle, and if the motion of the latter is confined to a limited region of the x-axis, e.g. between x_1 and x_2, then the average value of this quantity in the sense of classical mechanics, i.e. with respect to the time, can

be defined by the expression

$$\bar{f} = \frac{1}{T}\int_0^T f(x)\,dt \tag{24}$$

taken for a 'round trip' of the particle, T representing the duration of this round trip. The round trip can obviously be replaced by a one-way trip, since the motion must proceed in the same manner on the two halves of a round trip, with the sign of the velocity reversed. We can thus put

$$\bar{f} = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} f(x)\,dt,$$

where t_1 and t_2 denote the time of starting from the point x_1 and arriving at the point x_2 respectively. Replacing dt by dx/v, where v is a function of x determined by the equation $v^2 = \sqrt{\left\{\frac{2}{m}\{W-U(x)\}\right\}}$, we get

$$\bar{f} = \frac{1}{t_2-t_1}\int_{x_1}^{x_2}\frac{f(x)}{|v|}\,dx, \tag{24a}$$

or, if a 'round trip' is taken instead of a 'one-way' trip,

$$\bar{f} = \frac{1}{T}\oint\frac{f(x)}{v}\,dx,$$

the velocity v being taken with the same sign as dx (i.e. + when x is increasing from x_1 to x_2, and − when it is decreasing from x_2 to x_1).

Now the expression (24a) for $\bar{f}$ is identical with that obtained by means of the wave-mechanical definition of the average value of $f(x)$ according to the formula

$$\bar{f} = \int_{x_1}^{x_2} f(x)\psi\psi^*\,dx. \tag{24b}$$

if the function ψ is assumed to vanish outside the region (x_1, x_2) and is replaced by its approximate expression (23a) for this region. The normalization constant C must be determined by the condition $\int_{x_1}^{x_2}\psi\psi^*\,dx = 1$, that is,

$$C^2\int_{x_1}^{x_2}\frac{dx}{v} = C^2\int_{t_1}^{t_2} dt = C^2(t_2-t_1) = 1.$$

This agreement of the classical theory with the wave-mechanical theory must not be overestimated. As a matter of fact the function ψ does

not in general vanish outside the classically allowed region, but, as we have just seen, decreases there approximately as $e^{-2\pi|S^0|/h}$. According to the relation $v = \frac{1}{m}\frac{dS^0}{dx}$, we can put (dropping the term containing the time)

$$S^0 = m\int v\,dx = \int \sqrt{\{2m(W-U)\}}\,dx. \tag{25}$$

This formula applies just as well, i.e. with the same degree of approximation, to the points inside and outside the region (x_1, x_2). In the latter case, for a point $x > x_2$, we can put

$$|S^0(x)| = \int_{x_2}^{x} \sqrt{\{2m(U-W)\}}\,dx, \tag{25a}$$

and consequently

$$|\psi| = Ce^{-\frac{2\pi}{h}\int_{x_2}^{x}\sqrt{[2m(U-W)]}\,dx}. \tag{25b}$$

Thus, to the degree of approximation used, we should define the wave-mechanical average of $f(x)$ by the equation

$$\bar{f} = \int_{-\infty}^{+\infty} f(x)|\psi|^2\,dx$$

with

$$|\psi|^2 = \frac{C^2}{|v|} = C^2\Big/\sqrt{\Big\{\frac{2}{m}(W-U)\Big\}}$$

for $W \geqslant U$, i.e. for $x_1 \leqslant x \leqslant x_2$,

and

$$|\psi|^2 = C^2e^{-\frac{4\pi}{h}\int_{x_2}^{x}\sqrt{[2m(U-W)]}\,dx}$$

for $x > x_2$ and a similar expression for $x < x_1$. The constant C must be determined from the equation $\int_{-\infty}^{+\infty}|\psi|^2\,dx = 1$.

The difference between the classical and the wave-mechanical averages becomes particularly important when there are two or more classically allowed regions separated from one another by regions for which $W < U$. The latter, being permeable to the particle from the wave-mechanical point of view, do not actually separate but, on the contrary, connect the former regions.

The comparison of the classical 'time-average' with the wave-mechanical 'probable value' for the case of a three-dimensional motion is much more complicated than in the one-dimensional case and will be considered in the next section in connexion with the wave-mechanical interpretation of the quantum conditions. It must be remarked here that such averages or probable values have a meaning only when the motion is confined to a classically limited region, and that these limits

can be assigned *a priori* only in the case of a conservative motion, i.e. a motion with a given (constant) value of the energy W. Within the allowed region, limited by the surface $W-U=0$, the amplitude function A must satisfy the equation

$$\operatorname{div}(A^2\nabla s_0)=0,$$

which can be solved after the function s_0 has been determined from the Hamilton-Jacobi equation (17). It should be remembered that this equation, which represents another form of equation (17 a), expresses the law of the conservation of the copies of the particle, or of the probability of its location [cf. (19)].

Although there is in general no exact equivalence between the classical and the wave-mechanical average values, yet there are special cases when this equivalence turns out to be exact. An interesting case of this sort is provided by the so-called 'virial', i.e. by the quantity

$$V=\frac{\partial U}{\partial x}x+\frac{\partial U}{\partial y}y+\frac{\partial U}{\partial z}z,$$

which was introduced by Clausius in the kinetic theory of gases.

For a motion restricted to a limited region, the time average of this quantity $\overline{V}$ is connected with the time average of the kinetic energy by the relation

$$2\overline{T}=\overline{V}. \tag{26}$$

This is called the 'virial theorem'. It can be derived as follows: We multiply Newton's equations of motion

$$m_k\frac{d^2x_k}{dt^2}=-\frac{\partial U}{\partial x}, \text{ etc.},$$

by the corresponding coordinates and write

$$x_k\frac{d^2x_k}{dt^2}=\frac{d}{dt}\left(x_k\frac{dx_k}{dt}\right)-\left(\frac{dx_k}{dt}\right)^2.$$

Adding these transformed equations, we get

$$\frac{d}{dt}\sum_k m_k\left(x_k\frac{dx_k}{dt}+\dots\right)-\sum_k m_k\left[\left(\frac{dx_k}{dt}\right)^2+\dots\right]=-\sum_k\left(\frac{\partial U}{\partial x_k}x_k+\dots\right).$$

Formula (26) is then obtained by averaging with respect to the time and taking account of the fact that the mean value of

$$\frac{d}{dt}\sum m_k\left(x_k\frac{dx_k}{dt}+\dots\right)$$

vanishes. If we replace the kinetic energy T by the difference $W-U$ and assume that the potential energy is a homogeneous function of the

nth degree in the coordinates, formula (26) reduces to the form $2(\overline{W-U}) = n\overline{U}$ or

$$\overline{U} = \frac{2}{n+2} W. \tag{26a}$$

It can easily be shown that this relation remains exactly valid in wave mechanics if $\overline{U}$ is defined as the integral $\int U\psi\psi^* \, dV$ and ψ is defined as the exact solution of the corresponding Schrödinger equation. As an example we shall consider the simplest case of a one-dimensional wave-mechanical problem which is described by the equation

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(W-U)\psi = 0.$$

If we multiply this equation by $x\, d\psi^*/dx$ and the conjugate equation $\frac{d^2\psi^*}{dx^2} + \frac{8\pi^2 m}{h^2}(W-U)\psi^* = 0$ by $x\frac{d\psi}{dx}$ and add, we obtain

$$x\frac{d}{dx}\left(\frac{d\psi}{dx}\frac{d\psi^*}{dx}\right) + \frac{8\pi^2 m}{h^2} Wx\frac{d}{dx}(\psi\psi^*) - \frac{8\pi^2 m}{h^2} Ux\frac{d}{dx}(\psi\psi^*) = 0.$$

By partial integration with respect to x, taking into account the boundary conditions ($\psi = 0$ and $d\psi/dx = 0$ for $x = \pm\infty$), we get

$$-\int_{-\infty}^{+\infty} \frac{d\psi}{dx}\frac{d\psi^*}{dx}\, dx - \frac{8\pi^2 m}{h^2} W \int_{-\infty}^{+\infty} \psi\psi^* \, dx + \frac{8\pi^2 m}{h^2} \int_{-\infty}^{+\infty} \psi\psi^* \frac{d(Ux)}{dx}\, dx = 0,$$

or, since $\int_{-\infty}^{+\infty} \psi\psi^* \, dx = 1$ and $\int_{-\infty}^{+\infty} f\psi\psi^* \, dx = \bar{f}$,

$$\int_{-\infty}^{+\infty} \frac{d\psi}{dx}\frac{d\psi^*}{dx}\, dx + \frac{8\pi^2 m}{h^2}\left[W - \overline{\frac{d(Ux)}{dx}}\right] = 0.$$

Further, by multiplying the Schrödinger equation by ψ^*, we obtain

$$\int_{-\infty}^{+\infty} \psi^* \frac{d^2\psi}{dx^2}\, dx + \frac{8\pi^2 m}{h^2} \int_{-\infty}^{+\infty} (W-U)\psi\psi^* \, dx = 0,$$

i.e.

$$\int_{-\infty}^{+\infty} \psi^* \frac{d^2\psi}{dx^2}\, dx + \frac{8\pi^2 m}{h^2}(W-\overline{U}) = 0,$$

or, transforming the first term by partial integration,

$$-\int_{-\infty}^{+\infty} \frac{d\psi}{dx}\frac{d\psi^*}{dx}\, dx + \frac{8\pi^2 m}{h^2}(W-\overline{U}) = 0.$$

We have therefore $$W-\overline{U}+W-\frac{\overline{d(Ux)}}{dx}=0$$

or $$2(W-\overline{U})=x\overline{\frac{dU}{dx}}.$$

This is exactly formula (26) for the special case that we have considered.†

Another illustration of the connexion between the wave-mechanical and the classical theory is given by the similarity of the classical equations of motion,

$$m\frac{d^2x}{dt^2}=-\frac{\partial U}{\partial x}, \text{ etc.},$$

and the wave-mechanical relations

$$m\frac{d^2\bar{x}}{dt^2}=-\overline{\frac{\partial U}{\partial x}}, \text{ etc.}, \tag{27}$$

between the corresponding average (or probable) values of the quantities involved.

The relations (27) were found by P. Ehrenfest. They are usually referred to, in connexion with the propagation of a *wave packet*, as the equations of motion of the 'centre' or 'centroid' of the latter, that is, of the point with the coordinates

$$\bar{x}=\int x\psi\psi^*\,dV, \qquad \bar{y}=\int y\psi\psi^*\,dV, \qquad \bar{z}=\int z\psi\psi^*\,dV. \tag{27 a}$$

If the wave function ψ represents a wave packet formed by superposing waves with slightly different frequencies (i.e. motions with slightly different energies), the coordinates $\bar{x}$, $\bar{y}$, $\bar{z}$ are certain functions of the time (in the case of a stationary state where the dependence of ψ upon the time is specified by the factor $e^{-i2\pi\nu t}$ they reduce to constants), so that we can differentiate them with regard to the time. The corresponding quantities can be defined as the average values of the components of the velocity of the particle or its acceleration, etc.

We shall prove the relations (27) for the simplest case of a motion parallel to the x-axis (the proof can easily be extended to the case of three-dimensional motion). We have, by the definition of $\bar{x}$,

$$\frac{d\bar{x}}{dt}=\int_{-\infty}^{+\infty}x\frac{\partial(\psi\psi^*)}{\partial t}\,dx=\int_{-\infty}^{+\infty}x\left(\psi^*\frac{\partial\psi}{\partial t}+\psi\frac{\partial\psi^*}{\partial t}\right)dx,$$

since x and t are independent variables.

† The proof given is due to B. Finkelstein.

Now ψ and ψ^* satisfy the equations

$$\frac{\partial\psi}{\partial t} = \frac{i}{4\pi}\frac{h}{m}\left(\frac{\partial^2\psi}{\partial x^2} - \mu U\psi\right)$$

$$\frac{\partial\psi^*}{\partial t} = -\frac{i}{4\pi}\frac{h}{m}\left(\frac{\partial^2\psi^*}{\partial x^2} - \mu U\psi^*\right),$$

where $\mu = \dfrac{8\pi^2 m}{h^2}$. Hence

$$\frac{d\bar{x}}{dt} = \frac{ih}{4\pi m}\int_{-\infty}^{+\infty} x\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right)dx.$$

By partial integration, in conjunction with the fact that

$$\int_{-\infty}^{+\infty}\frac{df}{dx}\,dx = f(+\infty) - f(-\infty)$$

vanishes if the function f contains x or $d\psi/dx$ as a factor (since $\int_{-\infty}^{+\infty}\psi\psi^*\,dx$ must be finite and equal to 1), we obtain

$$\frac{d\bar{x}}{dt} = \frac{h}{4\pi m i}\int_{-\infty}^{+\infty}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right)dx. \tag{27 b}$$

This expression could be obtained directly from the relation $\dfrac{\partial(\psi\psi^*)}{\partial t} + \dfrac{\partial j}{\partial x} = 0$ (which is a special case of (19)) and the formula $j = \dfrac{h}{4\pi i m}\left(\psi^*\dfrac{\partial\psi}{\partial x} - \psi\dfrac{\partial\psi^*}{\partial x}\right)$ for the current density. Putting $j = \psi\psi^*\bar{v}(x)$, where $\bar{v}(x)$ is the average velocity at the point x, we can rewrite the preceding equation in the form

$$\frac{d\bar{x}}{dt} = \int_{-\infty}^{+\infty}\bar{v}(x)\psi\psi^*\,dx,$$

which agrees with the definition of $d\bar{x}/dt$ as the average value of the velocity of the particle irrespective of its position.

By differentiating (27 b) with respect to the time, we obtain

$$\frac{d^2\bar{x}}{dt^2} = -\frac{h^2}{16\pi^2 m^2}\int_{-\infty}^{+\infty}dx\left[\left(\frac{\partial^2\psi^*}{\partial x^2} - \mu U\psi^*\right)\frac{\partial\psi}{\partial x} - \psi^*\frac{\partial}{\partial x}\left(\frac{\partial^2\psi}{\partial x^2} - \mu U\psi\right) + \right.$$
$$\left. + \left(\frac{\partial^2\psi}{\partial x^2} - \mu U\psi\right)\frac{\partial\psi^*}{\partial x} - \psi\frac{\partial}{\partial x}\left(\frac{\partial^2\psi^*}{\partial x^2} - \mu U\psi^*\right)\right].$$

$$= -\frac{h^2}{8\pi^2 m^2}\int_{-\infty}^{+\infty}\left[\left(\frac{\partial^2\psi^*}{\partial x^2}-\mu U\psi^*\right)\frac{\partial\psi}{\partial x}+\left(\frac{\partial^2\psi}{\partial x^2}-\mu U\psi\right)\frac{\partial\psi^*}{\partial x}\right]dx$$

$$= -\frac{h^2}{8\pi^2 m^2}\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial x}\left(\frac{\partial\psi}{\partial x}\frac{\partial\psi^*}{\partial x}\right)-\mu U\frac{\partial}{\partial x}(\psi\psi^*)\right]dx$$

$$= -\frac{h^2\mu}{8\pi^2 m^2}\int_{-\infty}^{+\infty}\frac{\partial U}{\partial x}\psi\psi^*\,dx;$$

i.e.
$$m\frac{d^2\bar{x}}{dt^2} = -\frac{\overline{\partial U}}{\partial x},$$

where
$$\frac{\overline{\partial U}}{\partial x} = \int_{-\infty}^{+\infty}\frac{\partial U}{\partial x}\psi\psi^*\,dx$$

is the average (or probable) value of the force acting on the particle. It must be emphasized that this value refers not to the average (or probable) position of the particle, determined by the centre of the packet (otherwise this centre would move exactly according to the classical mechanics), but to all possible positions.

If the dimensions of the packet are very small (which means that the uncertainty in the estimation of the particle's velocity is very large) the motion of its centre closely follows classical motion. This, however, persists only for a very short time, for the packet will spread, the rate of this spreading being the larger the smaller its original dimensions (i.e. the larger the original uncertainty in the velocity).

5. Motion in a Limited Region; Quantum Conditions and Average Values

We shall now investigate the case of a (three-dimensional) motion restricted classically to a finite region of space (where $W-U>0$), and derive the 'quantization rules' characteristic of such a motion with the help of the approximate wave-mechanical theory based on the classical determination of the phase or action function $s(=S^0)$ by means of the Hamilton-Jacobi equation. A motion of this kind must obviously have a periodic or quasi-periodic character, so that the path described by the particle may fill up the whole region or pass many times in various directions through the same or nearly the same point (as, for instance, in the simple case of the oscillatory motion of a particle along a straight line). If the particle is replaced by a continuous assembly of its copies, a more complicated picture results, different copies passing simul-

taneously through the same point with velocities which are in general different both in regard to direction and (if the field of force varies with the time) in regard to magnitude. The latter must, of course, remain a single-valued function of the coordinates in the case of motion with a given (constant) value of the total energy W. The function $\phi = s/m$, which can be defined as the velocity potential, must, however, in this case (as well as in the general case of non-conservative motion) be a *multiple-valued function* of the coordinates. Considering the copy assembly as a kind of fluid, we can illustrate the case in question by the familiar type of fluid motion with *closed stream-lines*, each stream-line representing the path of all the particles situated on it. In the associated wave picture these closed paths of the separate particles or copies must be interpreted as *closed rays*.

Now a fluid motion of this type can be irrotational if, for instance, the fluid is flowing in a closed tube or around some closed tube. The velocity v of the particles, as a function of their coordinates, can then be represented as the gradient of a potential ϕ, provided the latter is defined as a multiple-valued function of the coordinates. In fact, taking the integral of the velocity along a line σ connecting two points P_1 and P_2, then, since the projection v_σ of v on the line element $d\sigma$ is, by definition, equal to $d\phi/d\sigma$, we get

$$\int_{P_1}^{P_2} v_\sigma \, d\sigma = \phi(P_2) - \phi(P_1).$$

If the line is closed, i.e. if the points P_1 and P_2 coincide, this integral should be equal to zero, irrespective of the shape of the line, unless we assume that for closed lines of certain type the potential ϕ may change after a 'round trip' by an amount $\Delta\phi$ equal to the value of the integral $\oint v_\sigma \, d\sigma$ taken along the corresponding closed line. If the latter coincides with a stream-line, the integral will certainly be different from zero, since along this line we must have $v_\sigma = |v|$.

Now it can easily be proved that in the case of irrotational motion the integral $\oint v_\sigma \, d\sigma$, which is called the 'circulation', will have *the same value for all closed lines of the same family*, i.e. of the same general type. In the case of a fluid flowing around a closed tube along closed stream-lines (Fig. 1), we must distinguish closed lines of two families: those which do not surround the tube, and those which do. For the former the circulation will be equal to zero, while for the latter it will have a certain value different from zero. This result follows from the transformation of the line integral $\oint v_\sigma \, d\sigma$, by means of Stokes[illegible],

into the integral $\oint (\text{curl}\, \mathbf{v})_n \, dS$ over any surface S limited by the line σ. In the case of the lines of the first family the surface S will be situated entirely within the fluid, so that the integral will vanish, since the motion is supposed to be irrotational ($\text{curl}\, \mathbf{v} = 0$). In the case of the lines of the second family the surface S will cut the tube around which the fluid is flowing. Since for points inside the tube the idea of velocity has no meaning, we can replace the surface S by another surface S' bounded by two closed lines of the second family. Stokes's formula applied to this surface which lies wholly within the fluid, and for which therefore the integral $\oint (\text{curl}\, \mathbf{v})_n \, dS$ vanishes, leads to the result that

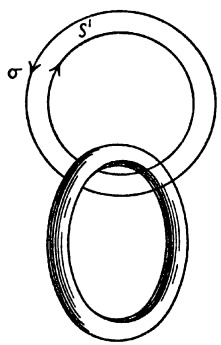

FIG. 1

the integral $\oint v_\sigma \, d\sigma$ taken over the double boundary of S' must vanish if the 'round trip' is made in opposite directions along the two constituent lines, whence it follows that the circulation will have the same value for both lines if the round trip is made in the same direction.

It may be mentioned that exactly similar results are met with in the theory of the magnetic field generated by a linear electric current. This field—outside the wire along which the current is flowing—is also irrotational, so that the magnetic field strength can be defined as the gradient of a certain magnetic potential. With every trip around the wire along any closed line (encircling this wire only once) this potential must change by a definite value, namely $4\pi i$, where i is the strength of the current.

The preceding results can be applied without substantial modification to the flow of the fictitious fluid represented by the copy assembly of a particle moving in a limited region. In the copy assembly, however, we must remember that different copies may be imagined to pass simultaneously through the same point in different directions. This is,

of course, impossible in the case of real particles. In particular, closed stream-lines may degenerate into 'double lines', i.e. unclosed lines along which the copies move first in one and then in the opposite direction (oscillatory motion).† The 'circulation' $\oint v_\sigma \, d\sigma$ for such a double line will not be equal to zero, but, on the contrary, will be equal to double the value of the integral $\int v_\sigma \, d\sigma$ for a one-way trip. As a result the velocity potential $\phi = s/m$, in addition to the multiplicity considered above, may acquire a duplicity of an entirely different character, corresponding to the possible presence at each point of two copies moving in opposite or, in general, in different directions.

Leaving aside this duplicity we see that, in the case of a particle confined to a finite region of space, the function s representing the mechanical action or the momentum-potential of the copies of this particle must—so long as the motion of these copies is supposed to be irrotational—be a multiple-valued function of the coordinates, i.e. it must change by a certain amount Δs for all closed lines (including double lines) of a certain family. It should be mentioned that 'round trips' along any of these lines have nothing to do with the actual motion, being performed not by definite copies (the latter need not move in closed lines), but by the process of linear integration referring to a *definite instant of time*. The change Δs of the function s for any such round trip is called a 'periodicity modulus' of s. From the point of view of the wave picture associated with the motion of the copy assembly of the particle these 'periodicity moduli' divided by the constant h represent the number of wave-lengths contained in the corresponding closed lines. In fact $ds/d\sigma = g_\sigma$ is the component of the momentum of the particle along the line-element $d\sigma$ and according to de Broglie's relation $d(s/h)/d\sigma = g_\sigma/h = k_\sigma$ must be equal to the corresponding component of the 'wave-number vector' $\mathbf{k} = \mathfrak{g}/h$ of the associated waves. The integral $\oint k_\sigma \, d\sigma = \Delta s/h$ may therefore be defined as the number of wave-lengths contained in the line σ, or, more exactly, as the number of wave-crests cut by this line, or still more exactly, as the difference between the number of waves cut by σ in the positive and in the negative direction (i.e. in the direction of propagation and in the opposite direction).

Now it is clear that in the case of motion corresponding to a definite energy, the wave system associated with it must be such that the number of waves cut by any closed line should be *integral*, corresponding

† The tube around which the fluid is supposed to flow degenerating into a ribbon with zero thickness.

to a change of the phase $\phi = 2\pi s/h$ by an integral multiple of 2π, a change which is irrelevant for the value of the wave function $\psi = Ae^{i\phi}$. In the contrary case the latter would also be a multiple-valued function of the coordinates, and would not represent a stationary system of standing waves (each standing wave being produced by the superposition of waves travelling in different directions), determined by the condition that the wave function ψ should vanish at or near the boundary of the region where the particle is supposed to move.

It thus follows from the condition of single-valuedness for the wave function ψ that *the 'periodicity moduli' of the 'action function' s must be integral multiples of h.*

This condition, which—it should be remembered—refers to the case of motion confined to a (classically) limited region, can easily be shown to be equivalent to the quantum conditions of the old quantum theory discovered by Bohr and by Sommerfeld.

For the general formulation of these quantum conditions, it is necessary, instead of the original rectangular coordinates x, y, z, to introduce new variables (generalized coordinates) q_1, q_2, q_3. If we succeed in so choosing these new variables that s assumes the form

$$s = \sum_{\alpha=1}^{f} s_\alpha(q_\alpha) \tag{28}$$

('separation variables'), then the quantum conditions run as follows:

$$\oint p_\alpha \, dq_\alpha = \oint \frac{ds_\alpha}{dq_\alpha} dq_\alpha = (\Delta s)_\alpha = n_\alpha h \quad (n_\alpha \text{ an integer}). \tag{28a}$$

Here the various $p_\alpha \left(= ds_\alpha/dq_\alpha\right)$ are the 'generalized momenta' and $(\Delta s)_\alpha$ are the 'principal moduli of periodicity' of the function s, i.e. those alterations of this function which correspond to a 'cyclic' change of one of the separation coordinates when the remaining two are kept fixed. By a 'cyclic' change of the coordinate q_α we mean an alteration such that the given particle returns to its original position and therefore the rectangular coordinates assume their original values. If the coordinate q_α has the character of an angle so that the rectangular coordinates are periodic functions of it, then the 'cyclic change' of q_α is simply the increase by the corresponding period Δq_α (for example, 2π). Otherwise it is an oscillation of q_α within certain limits determined by the nature of the field of force. The cyclic alterations of the individual separation coordinates in the actual motion of the system take place in periods of time Δt_α which are in general different from one another, so that the motion with regard to the time appears to be

non-periodic or *conditionally periodic*. This dependence of the variables q_α on the time plays no part in the 'quantizing' defined by formula (28 a).

The generalized momenta appearing in (28 a) can be defined, and indeed are usually defined, in a different way—namely, as the partial derivatives of the kinetic energy T, expressed as a function of the generalized coordinates and of the corresponding 'velocities' $dq_\alpha/dt = \dot{q}_\alpha$, with respect to the latter. The equivalence of both definitions is obvious in the case of rectangular coordinates, since $T = \frac{1}{2}m(v_x^2+v_y^2+v_z^2)$ and $g_x = \partial s/\partial x = \partial T/\partial v_x$, etc. If the coordinates are replaced by new (generalized) coordinates $q_\alpha(x, y, z)$, we have

$$\dot{q}_\alpha = \frac{\partial q_\alpha}{\partial x} v_x + \frac{\partial q_\alpha}{\partial y} v_y + \frac{\partial q_\alpha}{\partial z} v_z,$$

whence $\partial \dot{q}_\alpha/\partial v_x = \partial q_\alpha/\partial x$, etc. We thus get

$$\frac{\partial s}{\partial x} = \sum_{\alpha=1}^{3} \frac{\partial s}{\partial q_\alpha} \frac{\partial q_\alpha}{\partial x}, \qquad \frac{\partial T}{\partial v_x} = \sum_{\alpha=1}^{3} \frac{\partial T}{\partial \dot{q}_\alpha} \frac{\partial \dot{q}_\alpha}{\partial v_x} = \sum_{\alpha=1}^{3} \frac{\partial T}{\partial \dot{q}_\alpha} \frac{\partial q_\alpha}{\partial x},$$

and consequently,

$$\frac{\partial s}{\partial q_\alpha} = \frac{\partial T}{\partial \dot{q}_\alpha} = p_\alpha.$$

The formulation of the quantum conditions in the form (28 a) is sometimes possible in two or more different ways—if there exist several sets of 'separable' coordinates. Theoretically it is possible—in a single way at least—for any type of motion (restricted to a finite region). Practically, however, the 'separation coordinates' can be found only for simple types of motion (i.e. of the field of force). If the separation coordinates cannot be found, then the quantum conditions—in the sense of Bohr's theory—must be stated in the more general form indicated above, namely, that the moduli of periodicity of s with respect to any closed curve should be equal to an integral multiple of h (or to zero).

We shall now turn to the question of the relation between the wave-mechanical average or probable value of any function of the coordinates of the particle for a given quantized state of motion and the corresponding classical 'time average' of this function. The solution of this question depends upon the introduction of new coordinates of a still more general kind than those considered above in connexion with the formulation of the quantum conditions. These still more general coordinates are not directly expressible in terms of the original ones, but in terms of the original coordinates *and* the corresponding momenta, the new momenta being also functions of the old momenta and of the old coordinates.

Coordinate or rather coordinate-momenta transformations of this

type were introduced by Hamilton and are called contact or canonical transformations (the transformation considered above being a particular case of these transformations).

The theory of canonical transformations is based upon the preservation of the so-called 'canonical form' of the classical equations of motion. In the case of rectangular coordinates these canonical equations can be obtained directly from the usual equations of motion $m\,d^2x/dt^2 = -\partial U/\partial x$, etc., and have the form

$$\frac{dg_x}{dt} = -\frac{\partial H}{\partial x},\ldots; \qquad \frac{dx}{dt} = +\frac{\partial H}{\partial g_x},\ldots, \tag{29}$$

where

$$H = \frac{1}{2m}(g_x^2+g_y^2+g_z^2)+U \tag{29a}$$

is the total energy expressed as a function of the coordinates and momenta, and is usually denoted as the 'Hamiltonian function'. The equations (29) can be interpreted as referring to a particle moving not in ordinary space with the three coordinates x, y, z but in the six-dimensional *phase-space* (Part I, Chap. V) with the 'coordinates' x, y, z, g_x, g_y, g_z, the time derivatives of these coordinates representing the six components of the 'velocity' in phase-space and H being a function of the 'position' of the particle in the phase-space.†

For the sake of uniformity in notation we shall, in the following, instead of x, y, z write Q_1, Q_2, Q_3, and instead of g_x, g_y, g_z write P_1, P_2, P_3. The equations (29) then become

$$\frac{dP_\alpha}{dt} = -\frac{\partial H}{\partial Q_\alpha}, \qquad \frac{dQ_\alpha}{dt} = \frac{\partial H}{\partial P_\alpha}. \tag{29b}$$

We now introduce new coordinates Q'_1, Q'_2, Q'_3 determined by three equations of the form

$$Q'_\beta = Q'_\beta(Q_1, Q_2, Q_3) \text{ or } Q_\alpha = Q_\alpha(Q'_1, Q'_2, Q'_3) \qquad (\alpha, \beta = 1, 2, 3). \tag{30}$$

We then define the new momenta P'_1, P'_2, P'_3 by the formulae

$$P'_\beta = \frac{\partial s}{\partial Q'_\beta} = \sum_{\alpha=1}^{3} \frac{\partial s}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial Q'_\beta} = \sum_{\alpha=1}^{3} P_\alpha \frac{\partial Q_\alpha}{\partial Q'_\beta} \text{ or } P_\alpha = \sum_{\beta=1}^{3} P'_\beta \frac{\partial Q'_\beta}{\partial Q_\alpha}, \tag{30a}$$

which obviously do not assume a knowledge of the action function s. It can then easily be shown that these new coordinates and momenta satisfy a system of equations of the same form as (29b),

$$\frac{dP'_\beta}{dt} = -\frac{\partial H'}{\partial Q'_\beta}, \qquad \frac{dQ'_\beta}{dt} = \frac{\partial H'}{\partial P'_\beta} \qquad (\beta = 1, 2, 3), \tag{31}$$

† Instead of one particle one can consider a continuous assembly of its copies, distributed not in the ordinary space as before, but in the phase-space with a density depending in general upon the time.

where H' is the new Hamiltonian function which is obtained by replacing in the original function $H(Q, P)$ the old coordinates and momenta by the new, according to the formulae (30) and (30 a). The transformation defined by these formulae is called a 'point transformation'. As already mentioned, it is a special case of the *canonical* transformations. A canonical transformation (of the coordinates and momenta) is defined by the formulae

$$P_\alpha = \frac{\partial \Phi}{\partial Q_\alpha}, \qquad Q'_\beta = \frac{\partial \Phi}{\partial P'_\beta}, \tag{31 a}$$

where $\Phi(Q, P')$ is a completely arbitrary function of the *original* coordinates and the *new* momenta. If, in particular, we put

$$\Phi = \sum_{\beta=1}^{3} P'_\beta f_\beta(Q_1, Q_2, Q_3)$$

we obtain, by (31 a),

$$Q'_\beta = f_\beta(Q_1, Q_2, Q_3); \qquad P_\alpha = \sum_{\beta=1}^{3} P'_\beta \frac{\partial Q'_\beta}{\partial Q_\alpha},$$

which corresponds to the point transformation (30), (30 a).

The fact that the original canonical equations (29) are transformed by (31 a) into equations of the same canonical form (31) can be shown as follows:

We form the complete differential or rather the variation of the function Φ, corresponding to a virtual variation (completely independent of the actual motion) of the variables Q, P':

$$\delta\Phi = \sum_\alpha \frac{\partial \Phi}{\partial Q_\alpha} \delta Q_\alpha + \sum_\beta \frac{\partial \Phi}{\partial P'_\beta} \delta P'_\beta = \sum_\alpha P_\alpha \delta Q_\alpha + \sum_\beta Q'_\beta \delta P'_\beta,$$

and differentiate this expression with regard to the time. We also take the time derivative of Φ

$$\frac{d\Phi}{dt} = \sum_\alpha P_\alpha \frac{dQ_\alpha}{dt} + \sum_\beta Q'_\beta \frac{dP'_\beta}{dt},$$

and form its variation. By subtracting the expressions thus obtained, we get, remembering that δ and dt are commutative,

$$\sum_\alpha \left(\frac{dP_\alpha}{dt} \delta Q_\alpha - \frac{dQ_\alpha}{dt} \delta P_\alpha \right) = \sum_\beta \left(\frac{dP'_\beta}{dt} \delta Q'_\beta - \frac{dQ'_\beta}{dt} \delta P'_\beta \right).$$

Now by (29 b) we have

$$\sum_\alpha \left(\frac{dP_\alpha}{dt} \delta Q_\alpha - \frac{dQ_\alpha}{dt} \delta P_\alpha \right) = - \sum_\alpha \left(\frac{\partial H}{\partial Q_\alpha} \delta Q_\alpha + \frac{\partial H}{\partial P_\alpha} \delta P_\alpha \right) = -\delta H.$$

Hence, in virtue of $H(P, Q) = H'(P', Q')$,

we obtain

$$-\delta H' = -\sum_\beta \left(\frac{\partial H'}{\partial Q'_\beta}\delta Q'_\beta + \frac{\partial H'}{\partial P'_\beta}\delta P'_\beta\right) = \sum_\beta \left(\frac{dP'_\beta}{dt}\delta Q'_\beta - \frac{dQ'_\beta}{dt}\delta P'_\beta\right).$$

Since the variations δQ_β and $\delta P'_\beta$ are arbitrary, we can equate their coefficients. In this way we get equations (31).

Those canonical transformations, in which the transformed Hamiltonian H' depends only on the momenta P' and not on the coordinates Q', play a special role. Such coordinates are usually called *cyclic*. The equations (31) reduce in this case to

$$P'_\beta = \text{const.} \qquad \frac{dQ'_\beta}{dt} = \frac{\partial H'}{\partial P'_\beta} = \omega_\beta = \text{const.},$$

i.e.

$$Q'_\beta = \omega_\beta t + \phi_\beta.$$

If the transformation function Φ leading to cyclic coordinates is known, the mechanical problem can be regarded as solved, for the original coordinates and momenta are then expressed according to the equations (31 a) as functions of the time which, besides t, only contain constants P'_β, ω_β, and ϕ_β.

Now it follows from (31 a) that this special transformation function is just the action function s regarded as a function of Q_1, Q_2, Q_3 and of three arbitrary constants P'_1, P'_2, P'_3 which necessarily appear on solving the Hamilton-Jacobi equation (16) or (17) by which this function is defined. These constants of integration can be expressed in terms of the three principal *moduli* of periodicity of the action function $J_\alpha = (\Delta s)_\alpha$ with regard to a system of separable coordinates q_1, q_2, q_3 (which we need neither actually know nor consider in detail here). Replacing the original constants P'_α by their expressions in terms of J_1, J_2, J_3 we can write the transformation function Φ in the form $s(x, y, z; J_1, J_2, J_3)$ and *define the constants* J_α as the new momenta ($P'_\alpha = J_\alpha$). Considered from this point of view these constants are called the 'action variables' of the problem. The corresponding cyclic coordinates are called the 'angle variables'. We shall denote them by w_β ($= Q'_\beta$).

We have therefore

$$w_\beta = \omega_\beta t + \phi_\beta, \tag{32}$$

where according to the transformed canonical equations (31)

$$\omega_\beta = \frac{\partial H'}{\partial J_\beta} = \text{const.} \qquad (H' = W) \tag{32 a}$$

and

$$P_\alpha = \frac{\partial s}{\partial Q_\alpha}, \qquad w_\beta = \frac{\partial s}{\partial J_\beta}. \tag{32 b}$$

To ascertain the dependence of the old coordinates Q_α on the new coordinates w_β, we shall introduce for a moment as an intermediate link between them the separation coordinates q_1, q_2, q_3. Expressed as a function of the latter, the function s assumes the form

$$s = \sum_{\alpha=1}^{3} s_\alpha(q_\alpha, J_1, J_2, J_3).$$

To a cyclic alteration of the coordinate q_α there corresponds by (32 b) an alteration of the coordinate w_β by $\Delta_\alpha w_\beta = \Delta_\alpha \partial s_\alpha / \partial J_\beta$. We have therefore, because $\Delta_\alpha s_\beta = J_\alpha$ if $\alpha = \beta$, and $= 0$ if $\alpha \neq \beta$,

$$\Delta_\alpha w_\beta = \frac{\partial J_\alpha}{\partial J_\beta} = \begin{cases} 1 & (\alpha = \beta), \\ 0 & (\alpha \neq \beta). \end{cases}$$

These formulae show that when any angle variable w_β is increased by 1 and the remaining w's are maintained constant, which corresponds to the cyclic alteration of the separation coordinate q_β, i.e. to the return of the particle to the original position along a 'β-curve', then the action function s increases exactly by J_β.

From this it follows that the coordinates Q_β, and consequently the momenta P_β, are periodic functions of the angle coordinates with periods equal to 1. Each of them, as well as any function $f(Q_1, Q_2, Q_3)$ (or still more generally $f(Q, P)$), can be expressed in the form of a triple Fourier series

$$f = \sum_{k_1, k_2, k_3} f_{k_1, k_2, k_3} e^{i2\pi(k_1 w_1 + k_2 w_2 + k_3 w_3)}, \tag{33}$$

where k_1, k_2, k_3 are integers which can assume all values from $-\infty$ to $+\infty$, and f_{k_1, k_2, k_3} are certain expansion coefficients characteristic of the function f. If instead of the w_β we put their values obtained from (32), we get

$$f = \sum_{k_1, k_2, k_3} C_{k_1, k_2, k_3} e^{i2\pi(k_1 \omega_1 + k_2 \omega_2 + k_3 \omega_3)t}, \tag{33 a}$$

where the C_k are new expansion coefficients which we can regard as the amplitudes of various harmonic vibrations, while

$$\omega = k_1 \omega_1 + k_2 \omega_2 + k_3 \omega_3 \tag{33 b}$$

are the *frequencies* of these vibrations. The quantities ω_β, i.e. the velocities corresponding to the angle coordinates, represent therefore the *fundamental frequencies* of the motion.

We can now return to the problem of determining the time mean value of f. This problem can be solved at once by means of formula (33 a). Indeed, the required time mean value must obviously be equal to that amplitude coefficient in (33 a) for which the vibration frequency ω vanishes—or the sum of such coefficients if the equation $\omega = 0$ is satisfied by several different combinations of the numbers k_1, k_2, k_3.

This mean value can be represented on the one hand by the general formula $\bar{f} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f\,dt$. On the other hand it can be represented just as well by the formula

$$\bar{f} = \int_0^1\int_0^1\int_0^1 f\,dw_1\,dw_2\,dw_3 \tag{34}$$

which does not contain the time explicitly, the triple integration being extended over the 'period cube' in the coordinate space of the angle variables; f is given as a function of the angle variables by formula (33).

The expression (34) has the form of a 'statistical' mean value corresponding to an averaging over the various copies of the given particle distributed with a *constant density in the space of the angle coordinates* w_1, w_2, w_3. Its numerical agreement with the time mean value of f for a definite copy means that the curve described by the motion of such a copy fills up this space uniformly.†

We can now return from the angle coordinates to our original rectangular coordinates $Q_1 = x$, $Q_2 = y$, $Q_3 = z$. In view of the fact that the new momenta are constants, the old coordinates may be considered practically as functions of the new coordinates alone, and vice versa. We can thus transform the volume integral (34) according to the well-known theorem of Jacobi, and put

$$\bar{f} = \int fD\,dV, \tag{34a}$$

where $dV = dxdydz$ and

$$D = \begin{vmatrix} \frac{\partial w_1}{\partial x}, & \frac{\partial w_1}{\partial y}, & \frac{\partial w_1}{\partial z} \\ \frac{\partial w_2}{\partial x}, & \frac{\partial w_2}{\partial y}, & \frac{\partial w_2}{\partial z} \\ \frac{\partial w_3}{\partial x}, & \frac{\partial w_3}{\partial y}, & \frac{\partial w_3}{\partial z} \end{vmatrix}.$$

By (32b) this functional determinant can be written in the form

$$D = \begin{vmatrix} \frac{\partial^2 s}{\partial J_1 \partial x}, & \frac{\partial^2 s}{\partial J_1 \partial y}, & \frac{\partial^2 s}{\partial J_1 \partial z} \\ \frac{\partial^2 s}{\partial J_2 \partial x}, & \frac{\partial^2 s}{\partial J_2 \partial y}, & \frac{\partial^2 s}{\partial J_2 \partial z} \\ \frac{\partial^2 s}{\partial J_3 \partial x}, & \frac{\partial^2 s}{\partial J_3 \partial y}, & \frac{\partial^2 s}{\partial J_3 \partial z} \end{vmatrix}. \tag{34b}$$

† This condition is satisfied for non-degenerate motion, that is, motion for which the three fundamental frequencies ω_1, ω_1, ω_3 are not commensurable with each other.

The volume integration in (34a) must be extended over the whole region for which $W-U \geqslant 0$. We are thus brought to the conclusion that the relative probability that the particle will be found in the volume-element dV, as measured by the relative *duration* of its presence in this volume-element, is equal to D ($\int D\,dV = 1$). Comparing this result with the wave-mechanical average

$$\bar{f} = \int f\psi\psi^*\,dV,$$

we see that it will agree approximately with (34a) if $\psi\psi^* = D$. Now in the region $W-U \geqslant 0$ the function $s = S^0$ is real, so that the modulus of the function $\psi = Ae^{i2\pi s/h} = e^{i2\pi S^0/h+S'}$ must reduce to $A = e^{S'}$. It follows therefore that

$$A^2 = D.$$

It should be remembered that an *exact* agreement between the classical and the wave-mechanical mean value is out of the question—not only because of the approximative character of the preceding expression for ψ (with s determined from the Hamilton-Jacobi equation), but also because in the wave-mechanical case the integration must be extended over *all space* including the classically forbidden region. However, this region, although infinite, contributes in general only a finite and usually a small amount to the integral $\int f\psi\psi^*\,dV$ because of a very rapid decrease of the function $|\psi|^2$.

The relation $A^2 = D$ can of course be derived in a straightforward way by integrating the equation

$$\operatorname{div} A^2\nabla s = 0$$

[cf. (17a)], or the equation

$$\nabla^2 S^0 + 2\nabla S^0 \cdot \nabla S' = 0$$

to which (21b) is reduced in the case of conservative motion. This integration has been carried out (in the case of the second equation) by Van Vleck, who showed that A^2 must be proportional to the determinant

$$\begin{vmatrix} \frac{\partial^2 s}{\partial x\partial\alpha} & \frac{\partial^2 s}{\partial y\partial\alpha} & \frac{\partial^2 s}{\partial z\partial\alpha} \\ \frac{\partial^2 s}{\partial x\partial\beta} & \frac{\partial^2 s}{\partial y\partial\beta} & \frac{\partial^2 s}{\partial z\partial\beta} \\ \frac{\partial^2 s}{\partial x\partial\gamma} & \frac{\partial^2 s}{\partial y\partial\gamma} & \frac{\partial^2 s}{\partial z\partial\gamma} \end{vmatrix},$$

where α, β, γ are any three integration constants occurring in the expression of the function $s(x, y, z; \alpha, \beta, \gamma)$. This determinant is equal

to the product of D with the determinant $\dfrac{\partial(\alpha,\beta,\gamma)}{\partial(J_1,J_2,J_3)}$ which is a constant factor playing the role of a normalization constant.

In the special case of uni-dimensional motion the determinant (34 b) reduces to $\partial^2 s/\partial x \partial J$, whereas by direct integration we obtained, in this case, $A^2 = \dfrac{C^2}{v} = mC^2\Big/\dfrac{\partial s}{\partial x}$. Thus we must have

$$\frac{\partial^2 s}{\partial x \partial J} = mC^2\Big/\frac{\partial s}{\partial x},$$

that is,

$$\frac{\partial s}{\partial x}\frac{\partial}{\partial J}\left(\frac{\partial s}{\partial x}\right) = \frac{\partial}{\partial J}\frac{1}{2}\left(\frac{\partial s}{\partial x}\right)^2 = mC^2,$$

or since $\dfrac{1}{2m}\left(\dfrac{\partial s}{\partial x}\right)^2 = W-U$, we get $\dfrac{\partial}{\partial J}(W-U) = C^2$. This condition is actually fulfilled, for $\partial U/\partial J = 0$ and $\partial W/\partial J = \omega = 1/T$, where T is the period of motion [according to (32 a) with $W = H'$]. Hence we get $C^2 = 1/T$ in accordance with the simple theory developed in the preceding section.

II

OPERATORS

6. Operational Form of Schrödinger's Equation, and Operational Representation of Physical Quantities

The formal relation between classical mechanics and wave mechanics can be presented in another way which not only leads us to a deeper understanding of the theory but also to various important generalizations.

We can arrive at this relation by examining Schrödinger's equation (12) written in the form

$$D\psi = 0,$$

where D denotes the *operator*

$$D \equiv \frac{1}{2m}\left[\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}\right)^2+\left(\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)^2+\left(\frac{h}{2\pi i}\frac{\partial}{\partial z}\right)^2\right]+\frac{h}{2\pi i}\frac{\partial}{\partial t}+U.$$

This can be expressed in terms of the elementary differential operators

$$\frac{h}{2\pi i}\frac{\partial}{\partial x} = p_x, \qquad \frac{h}{2\pi i}\frac{\partial}{\partial y} = p_y, \qquad \frac{h}{2\pi i}\frac{\partial}{\partial z} = p_z, \qquad \frac{h}{2\pi i}\frac{\partial}{\partial t} = p_t \quad (35)$$

by the formula

$$D = \frac{1}{2m}(p_x^2+p_y^2+p_z^2)+p_t+U. \qquad (35\,a)$$

The equation $D\psi = 0$ thus reduces to the classical equation

$$T+U-W = 0$$

if we replace the operators p_x, p_y, p_z by the components of the momentum, and $-p_t$ by the total energy, i.e. if instead of (35) we put

$$p_x = g_x, \qquad p_y = g_y, \qquad p_z = g_z, \qquad p_t = -W \qquad (36)$$

and cancel the function ψ (considering it as a factor). Therefore the transition from classical mechanics to wave mechanics can formally be carried out as follows. In the 'classical' equation

$$\frac{1}{2m}(g_x^2+g_y^2+g_z^2)+U-W = 0, \qquad (36\,a)$$

which relates the components of the momentum and the total energy of a particle, we must replace these quantities by the elementary operators (35) and then multiply the Schrödinger operator D thus obtained by the wave function ψ on the right, where 'right multiplication' simply means applying the operator to the expression standing on its right.

The replacement of the energy W by the operator $-p_t = -\frac{h}{2\pi i}\frac{\partial}{\partial t}$ has

been made before, although in a somewhat different connexion, namely, in the transition from the wave equation

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(W-U)\psi = 0$$

for a conservative motion to the general equation

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}\left(-\frac{h}{2\pi i}\frac{\partial}{\partial t} - U\right)\psi = 0,$$

which applies to a motion of any kind. In the former case, since $\psi = \psi^0(x,y,z)e^{-i2\pi Wt/h}$, the operator p_t is actually equivalent to the energy in that it satisfies the equation $p_t\psi = -W\psi$, which we could write symbolically (dropping the function operated upon) in the form $p_t = -W$. A similar equivalence exists between the operators p_x, p_y, p_z and the components of the momentum g_x, g_y, g_z with respect to the wave function

$$\psi = \text{const.}\, e^{i2\pi(g_x x + g_y y + g_z z - Wt)/h},$$

representing the *free* motion of a particle with a velocity of specified magnitude and direction. As we know, the latter can be specified only in this particular case. In the general case the functions $p_x\psi$, $p_y\psi$, $p_z\psi$, $-p_t\psi$ are not equal to the products of the function ψ by constant numbers.

It is natural to associate this result with the fact that, in the general case, the components g_x, g_y, g_z of the momentum, as well as the energy W, cannot be defined as certain numbers since they do not have definite values, and to assume further that the operators p_x, p_y, p_z, $-p_t$ by which they are replaced in the transition from classical to wave mechanics must replace them in all wave-mechanical questions.

This principle is corroborated by the following considerations.

(1) If the wave function ψ can be approximated to by the expression $e^{i2\pi S/h}$ where S is the classical 'action', i.e. the momentum-potential determined by the Hamilton-Jacobi equation, then we have

$$p_x\psi = \frac{h}{2\pi i}\frac{\partial}{\partial x}e^{i2\pi S/h} = e^{i2\pi S/h}\frac{\partial S}{\partial x} = g_x\psi,$$

etc., so that in this approximation the operators p_x, p_y, p_z are actually equivalent to the components of the momentum g_x, g_y, g_z. This result still holds approximately if ψ is represented in the form $Ae^{i2\pi s/h}$ where s is the classical momentum-potential, for the partial derivatives of the amplitude A with regard to x, y, z (so far as the above approximation can be applied) are very small compared with the partial derivatives

of s/h, i.e. the components of the wave number (the wave-length being supposed to be very small).

(2) If the function ψ is 'quadratically integrable', i.e. if it can be normalized in such a way that the integral $\int \psi\psi^* \, dV$ is equal to 1, then the integrals

$$\int \psi^* p_x \psi \, dV, \qquad \int \psi^* p_y \psi \, dV, \qquad \int \psi^* {}_z p \psi \, dV$$

coincide with the average values of the components of the momentum as defined by the integrals

$$m \int j_x \, dV, \qquad m \int j_y \, dV, \qquad m \int j_z \, dV,$$

where $\mathbf{j} = \psi\psi^* \bar{\mathbf{v}}$ is the probability current density and $\bar{\mathbf{v}}$ is the average velocity introduced in the preceding chapter, §§ 2 and 3. We have in fact, according to the definition of j_x,

$$m \int j_x \, dV = \frac{h}{4\pi i} \int \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x} \right) dV.$$

Now by partial integration we get

$$\int \psi \frac{\partial \psi^*}{\partial x} \, dV = \int \frac{\partial}{\partial x} (\psi\psi^*) \, dV - \int \psi^* \frac{\partial \psi}{\partial x} \, dV = - \int \psi^* \frac{\partial \psi}{\partial x} dV,$$

since in order that $\int \psi\psi^* \, dV$ should have a finite value the function $\psi\psi^*$ must vanish at infinity rapidly enough to make the integral

$$\int \frac{\partial}{\partial x} (\psi\psi^*) \, dV = \int\int [\psi\psi^*]_{x=-\infty}^{x=+\infty} dydz$$

vanish too. Therefore

$$m \int j_x \, dV = \frac{h}{2\pi i} \int \psi^* \frac{\partial \psi}{\partial x} \, dV = \int \psi^* p_x \psi \, dV.$$

The preceding results can be extended to the more complicated operators, by which different classical quantities represented as certain functions of the coordinates and momenta $F(x, y, z; g_x, g_y, g_z)$ must be replaced, when g_x, g_y, g_z are replaced by the operators p_x, p_y, p_z. The simplest example of such a complicated operator is the operator $T = (p_x^2 + p_y^2 + p_z^2)/(2m)$ representing the kinetic energy. If the function ψ describes a motion with a given constant value of the total energy, i.e. if it satisfies the Schrödinger equation $(T + U - W)\psi = 0$, then we have $T\psi = (W - U)\psi$, where the 'operator' $(W - U)$ is a simple factor. The preceding equation expresses the fact that the kinetic energy (i.e. the magnitude of the classical velocity) is a definite function of the coordinates. The sum of the operator T and the potential energy U

represents the total energy of the particle and is usually called the energy operator, or the Hamiltonian operator, or simply the 'Hamiltonian'. Denoting this operator by H, we can write the preceding equation in the form $H\psi = W\psi$. It expresses the fact that the energy of the particle in the motion described by the function ψ has a definite value, namely, W. The general equation referring to a non-conservative motion can be written in the form

$$(H+p_t)\psi = 0. \tag{37}$$

It implies a certain relation between the two operators H and $-p_t$, both of which represent the energy W (when it exists)—the former in a specific way, including the properties of the particle (mass) and the character of the field of force in which it moves, and the latter in a perfectly general way independent of these characteristics.

Independently of the form of the operator $F(x, y, z; p_x, p_y, p_z)$, it can easily be shown that the result of applying it to the function ψ expressed in the approximate form $e^{i2\pi S/h}$ (or $Ae^{i2\pi s/h}$) is equal approximately to the product $F(x, y, z; g_x, g_y, g_z)\psi$. The same is true in the more general case of an operator containing the time t and the time derivative operator p_t. We have namely

$$F(x, y, z, t; p_x, p_y, p_z, p_t)\psi = F(x, y, z, t; g_x, g_y, g_z, -W)\psi,$$

if the energy W is defined as $-\partial S/\partial t$, in accordance with the Hamilton-Jacobi equation which gives $-\partial S/\partial t = (\nabla S)^2/2m+U = T+U$. The function $F\psi$ resulting from the application of the operator F to the *exact* wave function ψ can be represented as the *product* of the latter with a certain function F_C of the coordinates alone (and eventually of the time). The function $F_C = (F\psi)/\psi$ can be defined as the value of the quantity represented by the operator F at the corresponding point (and instant of time). This is precisely the way in which we have defined above the value of the kinetic energy in the case of a conservative motion. If, in particular, the ratio $(F\psi)/\psi$ is equal to a constant C, then the quantity represented by F is said to be a *constant of the motion*, its value C being independent of the position of the particle (and of the time). This case can be illustrated by applying the energy operator H to a function ψ which describes a conservative motion, or by applying any one of the operators p_x, p_y, p_z to the function ψ which describes a uniform rectilinear motion.

If the ratio $F_C = (F\psi)/\psi$ is not equal to a constant, then we can define the average or probable value of the quantity represented by

the operator F by means of the formula

$$\overline{F} = \int F_C \psi\psi^* \, dV$$

or

$$\overline{F} = \int \psi^* F \psi \, dV, \tag{38}$$

with the condition that

$$\int \psi\psi^* \, dV = 1. \tag{38 a}$$

This definition of an average value is a generalization of that already considered in the preceding chapter in connexion with quantities depending on the coordinates alone (such as the potential energy). Its physical significance has been tested above in the case of the fundamental operators p_x, p_y, p_z.

As a further illustration of the operational representation of physical quantities we shall consider the *angular momentum* of a particle, for instance, the angular momentum of an electron moving about a fixed nucleus (cf. Part I, § 14). In classical mechanics this quantity is defined as a vector with the components

$$yg_z - zg_y, \quad zg_x - xg_z, \quad xg_y - yg_x. \qquad = \vec{r} \times \vec{p}$$

We shall define it accordingly as a vector-operator M with the components

$$M_x = yp_z - zp_y, \qquad M_y = zp_x - xp_z, \qquad M_z = xp_y - yp_x,$$

or

$$\left.\begin{aligned} M_x &= \frac{h}{2\pi i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \qquad M_y = \frac{h}{2\pi i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \\ M_z &= \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \end{aligned}\right\}. \tag{39}$$

Transforming from rectangular coordinates to spherical coordinates by means of the formulae

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta,$$

we get

$$\frac{\partial\psi}{\partial r} = \frac{\partial\psi}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial\psi}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial\psi}{\partial z}\frac{\partial z}{\partial r},$$

i.e.

$$\begin{aligned} r\frac{\partial}{\partial r} &= r\sin\theta\cos\phi\frac{\partial}{\partial x} + r\sin\theta\sin\phi\frac{\partial}{\partial y} + r\cos\theta\frac{\partial}{\partial z} \\ &= x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}, \end{aligned}$$

and likewise

$$\frac{\partial}{\partial\phi} = -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\sin\theta\cos\phi\frac{\partial}{\partial y} = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

We have therefore

$$M_z = \frac{h}{2\pi i}\frac{\partial}{\partial\phi}. \tag{39 a}$$

Further, from (39) we get

$$\begin{aligned} M^2 &= M_x^2+M_y^2+M_z^2 \\ &= -\frac{h^2}{4\pi^2}\left[y^2\frac{\partial^2}{\partial z^2}+z^2\frac{\partial^2}{\partial y^2}-y\frac{\partial}{\partial z}\left(z\frac{\partial}{\partial y}\right)-z\frac{\partial}{\partial y}\left(y\frac{\partial}{\partial z}\right)+\ldots\right] \\ &= -\frac{h^2}{4\pi^2}\left[(y^2+z^2)\frac{\partial^2}{\partial x^2}-2yz\frac{\partial^2}{\partial y\partial z}-2x\frac{\partial}{\partial x}-\ldots\right] \\ &= -\frac{h^2}{4\pi^2}\left[(r^2-x^2)\frac{\partial^2}{\partial x^2}-2yz\frac{\partial^2}{\partial y\partial z}-2x\frac{\partial}{\partial x}-\ldots\right], \end{aligned}$$

where the terms denoted by ... are obtained from the given terms by cyclic permutation of the coordinates x, y, z. Because of the identity

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z}\right)^2 = x^2\frac{\partial^2}{\partial x^2}+\ldots+x\frac{\partial}{\partial x}+\ldots+2yz\frac{\partial^2}{\partial y\partial z}+\ldots,$$

or

$$x^2\frac{\partial^2}{\partial x^2}+\ldots+2yz\frac{\partial^2}{\partial y\partial z}+\ldots = \left(r\frac{\partial}{\partial r}\right)^2-r\frac{\partial}{\partial r},$$

we can write the previous expression in the form

$$\begin{aligned} M^2 &= -\frac{h^2}{4\pi^2}\left[r^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)-\left(r\frac{\partial}{\partial r}\right)^2-r\frac{\partial}{\partial r}\right] \\ &= -\frac{h^2}{4\pi^2}\left[r^2\nabla^2-r^2\frac{\partial^2}{\partial r^2}-2r\frac{\partial}{\partial r}\right]. \end{aligned}$$

Hence

$$\nabla^2 = -\frac{4\pi^2}{h^2}\frac{1}{r^2}M^2+\frac{1}{r^2}\left(r\frac{\partial}{\partial r}\right)^2+\frac{1}{r}\left(\frac{\partial}{\partial r}\right) = -\frac{4\pi^2}{h^2}\frac{M^2}{r^2}+\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r},$$

or putting

$$\nabla^2 = \frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\Omega^2,$$

where

$$\Omega^2 = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}$$

denotes the angular part of ∇^2, we get

$$M^2 = -\frac{h^2}{4\pi^2}\Omega^2. \tag{39 b}$$

By applying this operator and the operator (39 a) to the functions $\psi_{nlm} = F_{nl}(r)Y_{lm}(\theta,\phi)$, which specify the stationary states of a hydrogen-like atom, we get

$$M^2\psi_{nlm} = F_{nl}(r)M^2Y_{lm} = -\frac{h^2}{4\pi^2}F_{nl}\Omega^2Y_{lm},$$

and by the equation $\Omega^2Y_{lm}+l(l+1)Y_{lm} = 0$ we get

$$M^2\psi_{nlm} = \frac{h^2}{4\pi^2}l(l+1)\psi_{nlm}. \tag{40}$$

Since, further, the dependence of $Y_{lm}(\theta, \phi)$ upon ϕ is expressed by the factor $e^{im\phi}$,

$$M_z \psi_{nlm} = \frac{hm}{2\pi} \psi_{nlm}. \qquad (40\,a)$$

These relations show that the magnitude of the angular momentum as well as its direction are constants of the motion—just as in the classical theory of a particle moving in a central field of force. It should be mentioned that the character of the central field affects only the radial factor $F_{nl}(r)$ in the wave function ψ_{nlm}, the angular factor $Y_{lm}(\theta, \phi)$ being in all cases a spherical harmonic function. Therefore the above relations hold for the motion of a particle not only in a Coulomb field but in any central field of force. They show further that the quantum numbers l and m which have been introduced in Part I, § 14, as nodal numbers, characterizing the wave function ψ_{nlm} from a purely geometrical point of view, have also a dynamical meaning, one of them (l) determining the total magnitude of the angular momentum according to the relation $M^2 = l(l+1)h^2/4\pi^2$, and the other (m) determining the projection of the angular momentum upon the z-axis according to $M_z = mh/2\pi$. For this reason the numbers l and m will be called respectively the *angular* and the *axial* quantum numbers.† The constancy of the direction of the angular momentum is only proved indirectly by the relation (40 a) because the direction of the z-axis can be chosen arbitrarily, the functions ψ_{nlm} being so defined that the z-axis is the axis of the spherical harmonic functions $Y_{lm}(\theta, \phi) = P_{lm}(\theta) e^{im\phi}$. If we apply the operators M_x and M_y to these functions the result will not be similar to that obtained by applying the operator M_z because the functions $M_x \psi_{nlm}$ and $M_y \psi_{nlm}$ are not equal to multiples of ψ_{nlm}. Since we know that M_x and M_y also represent constants of the motion, we see that the condition $F\psi = \text{const.}\,\psi$ cannot be regarded as the general criterion for the constancy of the quantity represented by the operator F. It can easily be shown that the above failure of this equation to express the general condition of dynamical constancy is connected with *degeneracy*, i.e. with the fact that the functions ψ_{nlm} are not determined by the value of the energy W_n which, in fact, depends only on the 'principal' quantum number (n). Any linear combination of the n^2 functions ψ_{nlm}, which differ from one another by the values assigned to the numbers l and m, will also represent a stationary state belonging to the same value of the energy. This linear combination, i.e. the

† This seems preferable to the traditional denomination where l is referred to as the 'azimuthal' quantum number and m as the 'magnetic' quantum number.

coefficients C_{lm} in the sum $\sum_l \sum_m C_{lm}\psi_{nlm}$, can be so chosen that the resulting function ψ'_n will represent the same thing with respect to the x-axis as $\psi_{nl'm'}$ with respect to the z-axis. Applied to *this* function the operator M_x would be equivalent to multiplication by $m'h/2\pi$ according to the equation $M_x\psi'_n = (hm'/2\pi)\psi'_n$ which could be considered as a direct expression of the constancy of M_x. The function obtained by applying M_x to ψ_{nlm} can easily be shown to reduce to a linear combination $\sum_{m'=-l}^{+l} C_{m'}\psi_{nlm'}$ of the $2l+1$ functions ψ_{nlm} associated with the z-axis.

7. Characteristic Functions and Values of Operators; Operational Equations; Constants of the Motion

In general the equation $F\psi = \text{const.}\,\psi$ can only be satisfied by functions ψ of a special type which depend upon the nature of the operator F and are therefore called the *characteristic functions* of this operator ('Eigenfunktionen' of the German authors—often translated into English as 'proper functions'). The corresponding values of the constant factor are called the *characteristic values* of F. As an example we may take Schrödinger's equation $H\psi = W\psi$. In this equation the wave functions describing the stationary states of motion are the characteristic functions of the energy operator H, and the energy-levels W are its characteristic values. In the case of H, as well as in the case of any other operator, these values and the functions associated with them can form both a discrete and a continuous set. The characteristic functions are fully determined by an operator F for a one-dimensional problem, involving one coordinate only. In three-dimensional problems there remains in general a certain ambiguity in the choice of the functions ψ, as determined by a single equation of the type $F\psi = \text{const.}\,\psi$, an ambiguity which is known as 'degeneracy' if F is the energy operator H. Thus, for example, the operator $M_z = \frac{h}{2\pi i}\frac{\partial}{\partial\phi}$ specifies the corresponding characteristic functions only with regard to their dependence upon m, defining them as $\psi = f(r,\theta)e^{im\phi}$ where $f(r,\theta)$ is an arbitrary function of r and θ. The operator M^2 likewise determines the dependence of the characteristic functions on the angles θ, ϕ only, the equation $M^2\psi = \text{const.}\,\psi$ being satisfied by $\psi = f(r)Y_l(\theta,\phi)$ where $f(r)$ is an arbitrary function of r, and $Y_l(\theta,\phi)$ is an arbitrary spherical harmonic of order l, which can be expressed as a sum of $2l+1$ functions of the type $P_{lm}(\theta)e^{im\phi}$ with arbitrary coefficients.

Now we have also seen that Schrödinger's equation $H\psi = \text{const.}\,\psi$ in the case of a hydrogen-like atom has for each characteristic value of $H = W_n$ a solution of the form $\psi_n = f_n(r)Y(\theta, \phi)$, where $Y(\theta, \phi)$ is a sum of n^2 spherical harmonic functions of the type $P_{lm}(\theta)e^{im\phi}$ with arbitrary coefficients ($l = 0, 1, ..., n-1$; $m = -l, ..., +l$). We cannot therefore completely specify the functions ψ_{nlm} describing the stationary states of a hydrogen atom by taking one of the three equations

$$H\psi = \text{const.}\,\psi, \qquad M^2\psi = \text{const.}\,\psi, \qquad M_z\psi = \text{const.}\,\psi, \tag{41}$$

but only by taking all three equations together. The functions ψ_{nlm} then appear as the 'simultaneous characteristic functions' of the operators H, M^2, and M_z, each of these functions belonging to a 'triplet' of characteristic values W_n, $(M^2)_l = l(l+1)h^2/4\pi^2$, and $(M_z)_m = mh/2\pi$.

Another simple example of this relationship is provided by the operators p_x, p_y, p_z. The characteristic functions of these operators are obviously $f_1(y, z)e^{i2\pi g_x x/h}$, $f_2(z, x)e^{i2\pi g_y y/h}$, $f_3(x, y)e^{i2\pi g_z z/h}$; f_1, f_2, f_3 being arbitrary functions of the corresponding arguments. Taken together the three equations

$$p_x\psi = g_x\psi, \qquad p_y\psi = g_y\psi, \qquad p_z\psi = g_z\psi, \tag{41 a}$$

where g_x, g_y, g_z are constants, specify unambiguously the function

$$\psi = \text{const.}\, e^{i2\pi(g_x x + g_y y + g_z z)/h}, \tag{41 b}$$

which describes the uniform rectilinear motion of a particle with the momentum components g_x, g_y, g_z, and which is a *particular* solution of Schrödinger's equation $H\psi = W\psi$ with $H = (p_x^2+p_y^2+p_z^2)/2m$, i.e. with $U = 0$, corresponding to free motion.

It should be mentioned that the expression (41 b) for ψ is still incomplete (as well as the expression $\psi = f_n(r)Y_{lm}(\theta, \phi)$ for the hydrogen-like atom functions) inasmuch as it does not contain the *time*. The latter can be introduced by the additional relation

$$-p_t\psi = W\psi,$$

giving $\psi \sim e^{-i2\pi Wt/h}$. The constant W is, however, *not independent*, but is connected with g_x, g_y, g_z by the relation $W = (g_x^2+g_y^2+g_z^2)/2m$.

If F is an ordinary function of the coordinates (or of the time too) which does not contain the elementary differential operators p_x, p_y, p_z, then the equation $F\psi = \text{const.}\,\psi$ has no solutions of the ordinary continuous type. The only possible solutions—except the trivial one $\psi = 0$—are those for which the function ψ is different from zero on the surface $F = \text{const.}$ and vanishes outside this surface (which can be displaced by varying arbitrarily the value of the constant).

Another interesting case is provided by operators which satisfy the equation $F\psi = C\psi$ *identically*, i.e. irrespective of the choice of the function ψ, and therefore do not determine this function at all. $F = p_x x - x p_x$ is the simplest example of such an operator. Applying it to some function ψ, we get

$$F\psi = \frac{h}{2\pi i}\left[\frac{\partial}{\partial x}(x\psi) - x\frac{\partial}{\partial x}\psi\right] = \frac{h}{2\pi i}\psi.$$

Thus we see that this operator has one *single* characteristic value $C = h/2\pi i$ with which *any* function can be associated as a 'characteristic function'. The preceding equation can be written symbolically in the form

$$p_x x - x p_x = \frac{h}{2\pi i}, \tag{42}$$

which is obtained by omitting the arbitrary function ψ to which the left- and right-hand sides of this equation must be applied. We have, of course, similar equations for the two other coordinates and the corresponding components of the momentum-operator: $p_y y - y p_y = h/2\pi i$ and $p_z z - z p_z = h/2\pi i$. In addition we have the 'operational' equations $p_x y - y p_x = 0$ or $p_x y = y p_x$, etc., which express the fact that the order in which the operators p_x and y are applied to any function $\psi(x, y, z)$ is immaterial (since x and y are independent variables). The equations $p_x p_y - p_y p_x = 0$ are quite similar to the equations $xy - yx = 0$ expressing the commutative law of ordinary multiplication. Two operators F and G which, when applied successively in the order F, G to any function ψ give the same result as when applied in the opposite order G, F, are said to be *commutable*. This property is expressed symbolically by the operational equation

$$FG = GF, \tag{42 a}$$

which means that the ordinary equation

$$FG\psi = GF\psi$$

is satisfied *identically*, i.e. for any function ψ.

In general, the fact that the equation $A\psi = B\psi$ is satisfied *identically* with respect to the function ψ, A and B being two outwardly different operators, is expressed symbolically by the equation $A = B$. We shall now give a few examples of such operational equations.

Let us consider first of all the operator $F = p_x f - f p_x$ where $f(x, y, z)$ is an arbitrary (continuous) function of the coordinates. Applying it to an arbitrary function ψ, we get

$$F\psi = \frac{h}{2\pi i}\left[\frac{\partial}{\partial x}(f\psi) - f\frac{\partial \psi}{\partial x}\right] = \frac{h}{2\pi i}\frac{\partial f}{\partial x}\psi,$$

so that
$$p_x f - f p_x = \frac{h}{2\pi i}\frac{\partial f}{\partial x}, \tag{43}$$

which means that the operator $p_x f - f p_x$ is equivalent to the multiplier $\frac{h}{2\pi i}\frac{\partial f}{\partial x}$.

The preceding equation is often written in the form
$$\frac{\partial f}{\partial x} = [p_x, f], \tag{43a}$$

where the bracket expression on the right side is defined by
$$[p_x, f] = \frac{2\pi i}{h}(p_x f - f p_x). \tag{43b}$$

If, in the above definition of F, we replace f by x and p_x by p_x^n [which means differentiation of the nth order with regard to x, combined with a multiplication by $(h/2\pi i)^n$], we get
$$F\psi = \left(\frac{h}{2\pi i}\right)^n \left[\frac{\partial^n}{\partial x^n}(x\psi) - x\frac{\partial^n}{\partial x^n}\psi\right] = \left(\frac{h}{2\pi i}\right)^n n\frac{\partial^{n-1}}{\partial x^{n-1}}\psi = \frac{h}{2\pi i} n p_x^{n-1}\psi,$$

so that
$$p_x^n x - x p_x^n = \frac{h}{2\pi i} n p_x^{n-1}, \tag{44}$$

which can be rewritten symbolically in the form
$$x p_x^n - p_x^n x = -\frac{h}{2\pi i}\frac{\partial}{\partial p_x} p_x^n.$$

This formula can easily be generalized for any operator expressible as the sum of terms $a_n p_x^n$ with coefficients a_n which do not depend upon the coordinate x. Denoting this operator by $f(p_x, p_y, p_z; y, z)$, we get
$$xf - fx = -\frac{h}{2\pi i}\frac{\partial f}{\partial p_x}, \tag{44a}$$

an equation very similar to (43) with x playing the role of $-p_x$, and p_x the role of x. Putting
$$[x, f] = \frac{2\pi i}{h}(xf - fx) \tag{44b}$$

we can consider the equation
$$\frac{\partial f}{\partial p_x} = -[x, f] \tag{44c}$$

as the general definition of the operator $\partial/\partial p_x$. We shall write in general
$$[F, G] = \frac{2\pi i}{h}(FG - GF), \tag{45}$$

this 'bracket expression' introduced by Dirac as the quantum analogue of the Poisson brackets vanishing if the operators F and G commute with one another.

It should be noticed that an operational equation $A = B$ expresses the *identity* of the physical quantities represented by the operators A and B; the existence of such equations indicates that the same physical quantity can be represented in wave mechanics in a number of apparently different ways.

Another interesting and important illustration of operational equations is provided by the representation of the *angular momentum* of a particle.

From the definition (39) it follows that

$$M_x^2 = (yp_z - zp_y)^2 = (yp_z)^2 - (yp_z)(zp_y) - (zp_y)(yp_z) + (zp_y)^2$$
$$= y^2p_z^2 + z^2p_y^2 - yp_yp_zz - zp_zp_yy,$$

since p_y commutes with z and p_z, and p_z commutes with y and p_y. Taking into account the relations $p_zz = zp_z + h/2\pi i$ and $p_yy = yp_y + h/2\pi i$, we get

$$M_x^2 = y^2p_z^2 + z^2p_y^2 - 2yzp_yp_z - \frac{h}{2\pi i}(yp_y + zp_z),$$

whence the formula (39 b) can easily be obtained. We have in addition

$$M_xM_y = (yp_z - zp_y)(zp_x - xp_z) = yp_zzp_x - zp_yzp_x - yp_zxp_z + zp_yxp_z$$
$$= yp_xp_zz - z^2p_yp_x - yxp_z^2 + zxp_yp_z,$$

whence

$$M_xM_y - M_yM_x = yp_xp_zz + zxp_yp_z - xp_yp_zz - zyp_xp_z,$$
$$= (yp_x - xp_y)(p_zz - zp_z) = \frac{h}{2\pi i}(yp_x - xp_y) = -\frac{h}{2\pi i}M_z.$$

Thus, according to (45),

$$[M_x, M_y] = -M_z. \tag{45 a}$$

In a similar way we can derive the relations $[M_y, M_z] = -M_x$ and $[M_z, M_x] = -M_y$, which can also be obtained from (45 a) by a cyclic permutation of the indices x, y, z. These three relations can be replaced by the symbolic vector equation

$$\mathbf{M} \times \mathbf{M} = -\frac{h}{2\pi i}\mathbf{M}, \tag{45 b}$$

where $\mathbf{A} \times \mathbf{B}$ is defined in the usual way as the vector product of $\mathbf{A}$ and $\mathbf{B}$.

Interesting results are obtained by calculating the bracket expressions for the components of the vector $\mathbf{M}$ on the one hand, and the components of the vector $\mathbf{r}(x, y, z)$ or $\mathbf{p}(p_x, p_y, p_z)$ on the other. We shall not go into these calculations (which can easily be carried out by

the reader) but shall merely notice the following results:

$$[p^2, \mathbf{M}] = 0, \qquad [p^2, M^2] = 0, \tag{46}$$

where $p^2 = p_x^2+p_y^2+p_z^2$, the first of these equations being equivalent to the three equations $[p^2, M_x] = 0$, $[p^2, M_y] = 0$, $[p^2, M_z] = 0$. These equations express the fact that the angular momentum of a particle commutes with its kinetic energy $T = p^2/2m$ (more exactly we should speak of the *operators* representing the angular momentum and the kinetic energy). If the potential energy U is a function of the distance $r = \sqrt{\{x^2+y^2+z^2\}}$ alone (which corresponds to a central field of force), then we also have

$$[U, \mathbf{M}] = 0, \qquad [U, M^2] = 0, \tag{46 a}$$

and consequently

$$[H, \mathbf{M}] = 0, \qquad [H, M^2] = 0, \tag{46 b}$$

where $H = p^2/2m+U$ is the Hamiltonian operator representing the total energy of the particle.

The relations (46 b) can be obtained very simply by using *polar* coordinates to represent H and $\mathbf{M}$. Then

$$H = \frac{1}{2m}\left(\frac{h}{2\pi i}\right)^2\left[\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\Omega^2\right]+U(r),$$

$$M_z = \frac{h}{2\pi i}\frac{\partial}{\partial \phi}, \qquad M^2 = -\frac{h^2}{4\pi^2}\Omega^2,$$

and so

$$[H, M_z] = \frac{1}{2m}\left(\frac{h}{2\pi i}\right)^3\frac{1}{r^2}\left[\Omega^2, \frac{\partial}{\partial \phi}\right], \qquad [H, M^2] = \frac{1}{2m}\left(\frac{h}{2\pi i}\right)^4\frac{1}{r^2}[\Omega^2, \Omega^2],$$

both bracket expressions $[\Omega^2, \partial/\partial\phi]$ and $[\Omega^2, \Omega^2]$ obviously vanishing.†

The equations (46 b) must be naturally related to the fact that $\mathbf{M}$ and M^2 represent quantities which are constants of the motion (in the case of a radially symmetrical field of force). An equation of the type

$$[H, F] = 0, \tag{47}$$

i.e. the commutability of an operator F with the energy operator H, can actually be considered as the most *general expression of the fact that F represents a constant of the motion* determined by the operator H, i.e. by Schrödinger's equation $H\psi = W\psi$.

In fact, applying the operator F to both sides of this equation, we have $FH\psi = WF\psi$ or, if $HF = FH$, we obtain $H(F\psi) = W(F\psi)$. This shows that the function $F\psi$ satisfies the same equation as the function

† In order to obtain (46 a) without the use of polar coordinates we need only notice that $[U, M_z] = [U, yp_x - zp_y] = y[U, p_x] - z[U, p_y] = z\dfrac{\partial U}{\partial y} - y\dfrac{\partial U}{\partial z}$ according to (43 a).